

SKRIPSI

**IMPLEMENTASI SISTEM REKOMENDASI GENRE
FILM DAN MUSIK BERBASIS EMOSI PADA TEKS
PENGGUNA DENGAN METODE CONTENT-
BASED FILTERING.**

MUHAMMAD IQMAL BASORI
NPM 21081010297

DOSEN PEMBIMBING
Hendra Maulana S.Kom M.Kom

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2024**

BAB 1 PENDAHULUAN

1.1 Latar Belakang:

Di era digital, kemajuan teknologi telah mendorong pengembangan sistem yang semakin sesuai dengan kebutuhan dan preferensi pengguna. Salah satu inovasi yang semakin berkembang adalah sistem rekomendasi, yang bertujuan untuk memberikan saran personalisasi berdasarkan pola perilaku dan karakteristik pengguna. Sistem ini sangat penting untuk berbagai industri, seperti e-commerce, hiburan, dan layanan streaming, karena membantu pengguna menemukan konten yang relevan daripada mencari konten yang tidak relevan.

Kebanyakan sistem rekomendasi konvensional, bagaimanapun, hanya menggunakan data historis pengguna, seperti riwayat pencarian atau interaksi, tanpa memperhitungkan faktor emosional yang dapat memengaruhi preferensi pengguna terhadap konten tertentu. Dan sistem rekomendasi tradisional biasanya bergantung pada data historis seperti riwayat pencarian, interaksi pengguna, atau ulasan sebelumnya. Ini adalah masalah utama. Metode ini seringkali mengabaikan aspek kontekstual dan emosional yang memengaruhi preferensi pengguna dalam memilih konten, terutama untuk hiburan seperti film dan musik. Sebagai contoh, sistem mungkin merekomendasikan genre yang relevan berdasarkan riwayat pengguna, tetapi tidak mempertimbangkan bahwa preferensi pengguna dapat berubah secara dinamis sesuai dengan suasana hati atau emosi mereka saat itu. Akibatnya, rekomendasi yang dihasilkan cenderung kurang relevan.

Penelitian tentang sistem rekomendasi musik berbasis emosi terutama melalui deteksi ekspresi wajah telah berkembang pesat. Studi sebelumnya, seperti yang dilakukan oleh Mammadli et al. (2022) menunjukkan bahwa model berbasis Convolutional Neural Network (CNN) dapat mengidentifikasi emosi pengguna melalui analisis citra wajah. Meskipun metode ini cukup berhasil, karena ia tidak mencakup semua kategori emosi.

Fokus penelitian adalah membangun sistem rekomendasi berbasis emosi yang diidentifikasi dari teks yang ditulis oleh pengguna. Emosi yang diidentifikasi digunakan untuk memberikan rekomendasi tentang genre film dan musik yang sesuai. Emosi yang akan dideteksi oleh sistem rekomendasi yaitu **Marah, Takut, Bahagia, Sedih, Cinta**. Sistem ini menggunakan pendekatan

Content-Based Filtering yaitu pendekatan yang memanfaatkan karakteristik tertentu dari konten. Algoritma yang Digunakan yaitu:

1. Cosine Similarity

Mengukur kesamaan antara dua vektor berdasarkan sudut di antara mereka. Cocok untuk menghitung kesamaan antara profil pengguna (vektor fitur) dan item.

2. Euclidean Distance

Mengukur jarak antara dua vektor dalam ruang fitur. Item dengan jarak terpendek lebih mirip dengan profil pengguna.

3. K-Nearest Neighbors (KNN)

Algoritma berbasis instance untuk menemukan item yang paling mirip dengan item favorit pengguna.

Metode berikut memungkinkan sistem untuk memahami fitur lagu, seperti lirik, melodi, dan tempo, dan kemudian menghubungkannya dengan emosi pengguna yang akan dideteksi. Dengan demikian, pengguna akan memiliki pengalaman musik dan film yang sesuai dan relevan dengan suasana hati mereka.

1.2 Rumusan Masalah:

1. Bagaimana mendesain sistem rekomendasi yang akan dibangun dengan metode Content-Based Filtering untuk merekomendasikan sistem berbasis emosi ke pengguna?
2. Seberapa pengaruh pengembangan sistem rekomendasi berbasis emosi berdampak besar pada kebutuhan emosi pengguna?
3. Bagaimana mengintegrasikan pendekatan *content-based filtering* dari ketiga algoritma Cosine Similarity, Euclidean Distance, K-Nearest Neighbors (KNN) untuk meningkatkan akurasi dan relevansi rekomendasi sistem?
4. Bagaimana menciptakan model rekomendasi yang mendukung keberagaman genre musik dan film beserta menghasilkan saran musik dan film yang akan dituju secara random?
5. Seberapa akurat Content-Based Filtering untuk rekomendasi sistem yang sedang dikembangkan?

1.3 Tujuan Penelitian:

1. Mengembangkan sistem rekomendasi dengan metode Hybrid Filtering untuk deteksi emosi pengguna.
2. Mengukur seberapa besar pengaruh dalam pengembangan sistem rekomendasi pada kebutuhan rekomendasi emosi pengguna.
3. Mengintegrasikan pendekatan *content-based filtering* dari ketiga algoritma Cosine Similarity, Euclidean Distance, K-Nearest Neighbors (KNN) untuk meningkatkan akurasi dan relevansi rekomendasi sistem.
4. Merancang model rekomendasi musik dan film yang mendukung keberagaman genre musik dan film beserta menghasilkan saran musik dan film yang akan dituju secara random.
5. Mengukur seberapa akurat Content-Based Filtering untuk rekomendasi sistem.

1.4 Manfaat Penelitian:

1. Membantu pengguna dalam memilih film yang sesuai dengan suasana hati kebutuhan emosional mereka.
2. Memberikan kontribusi pada pengembangan ilmu pengetahuan di bidang sistem rekomendasi genre film dan musik, khususnya yang berbasis emosi.
3. Membantu platform hiburan film dan musik untuk meningkatkan pengalaman pengguna dengan rekomendasi yang lebih akurat, personal, dan relevan.

BAB 2 TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Penelitian tentang sistem rekomendasi berbasis emosi telah berkembang pesat dengan pendekatan yang menggabungkan teknologi kecerdasan buatan, pemrosesan bahasa alami (NLP), dan pembelajaran mesin. Berikut adalah beberapa penelitiannya:

2.1.1 Sistem Rekomendasi Berbasis Emosi

Mammadli et al. (2022) memanfaatkan model *Convolutional Neural Network* (CNN) untuk menganalisis emosi pengguna melalui deteksi ekspresi wajah. Hasil penelitian menunjukkan bahwa algoritma ini sangat efektif untuk mendeteksi emosi seperti bahagia, sedih, dan marah dengan akurasi tinggi. Namun, keterbatasannya adalah ketergantungan pada input visual, sehingga sulit diterapkan pada pengguna yang tidak ingin memberikan data wajah mereka. Ho et al. (2024) mengembangkan *Emotion-Based Movie Recommender System* (E-MRS) dengan menggabungkan *Content-Based Filtering* (CBF) dan *Collaborative Filtering* (CF). Sistem ini menggunakan emosi sebagai variabel utama dalam menentukan rekomendasi, mengatasi masalah perubahan preferensi pengguna secara dinamis.

2.1.2 Analisis Emosi dari Teks

Tripathi et al. (2018) menyoroti penggunaan teknik Term Frequency-Inverse Document Frequency (TF-IDF) sebagai metode representasi data untuk menangkap signifikansi relatif kata dalam teks tertentu. Representasi ini mempermudah algoritma pembelajaran mesin seperti Support Vector Machine (SVM) atau Conditional Random Fields (CRF) untuk mengklasifikasikan emosi seperti marah, sedih, bahagia, atau takut. Liu et al. (2003) memperkenalkan Common-Sense Affect Knowledge Base, yang menekankan hubungan semantik antara kata-kata dan emosi yang dihasilkan. Penelitian ini menggabungkan pendekatan berbasis aturan dan lexical affinity, menghasilkan sistem yang lebih peka terhadap nuansa linguistik seperti idiom dan metafora.

Teks pendek seperti tweet sering digunakan karena strukturnya yang sederhana dan langsung mencerminkan emosi pengguna. Aman dan Szpakowicz (2007) menggunakan hashtag seperti #happy dan #angry untuk memberikan label supervisi pada tweet, memungkinkan pelatihan model berbasis supervised learning dengan hasil yang akurat. Untuk teks panjang seperti cerita anak-anak, Alm et al. (2008) menganalisis emosi dengan membagi cerita menjadi kalimat

dan memberikan anotasi per kalimat. Penelitian ini menunjukkan bahwa struktur narasi yang kompleks dapat menangkap perubahan emosi secara bertahap, yang relevan dalam konteks seperti artikel atau ulasan panjang.

Meski metode berbasis teks telah menunjukkan hasil yang menjanjikan, beberapa tantangan utama tetap ada:

- Ambiguitas Semantik: Kata atau frasa yang sama dapat mengekspresikan emosi yang berbeda tergantung pada konteksnya.
- Kombinasi Emosi: Teks sering mencerminkan lebih dari satu emosi pada saat yang sama, yang mempersulit klasifikasi berbasis kategori tunggal.

2.1.3 Pendekatan Hybrid dalam Sistem Rekomendasi

Saifudin et al. (2023) melakukan tinjauan sistematis terhadap berbagai pendekatan hybrid dalam sistem rekomendasi. Penelitian mereka menunjukkan bahwa kombinasi antara *Content-Based Filtering* (CBF) dan *Collaborative Filtering* (CF) dapat secara signifikan mengatasi masalah sparsity dan cold start. Misalnya, pendekatan *model-based* dalam CF memungkinkan penggunaan algoritma seperti Singular Value Decomposition (SVD) untuk mengurangi dimensi data, yang menghasilkan performa lebih baik dibandingkan pendekatan tradisional. Bang et al. (2024) menggunakan hybrid filtering untuk sistem rekomendasi musik berbasis emosi dengan mengintegrasikan analisis data teks dan visual. Sistem ini berhasil menciptakan pengalaman personalisasi dengan akurasi validasi sebesar 97% menggunakan algoritma CNN pada dataset FER-2013.

2.1.4 Aplikasi pada Rekomendasi Musik dan Film

Patra et al. (2015) menyoroti pentingnya analisis lirik lagu untuk mengidentifikasi emosi pengguna. Dengan menggunakan kombinasi sentiment lexicons dan analisis n-grams, sistem ini berhasil memetakan suasana hati pengguna dengan lirik lagu untuk mencocokkan genre musik seperti klasik, jazz, atau pop.

Bang et al. (2024) menggunakan model CNN untuk mengenali ekspresi wajah pengguna dan menghubungkannya dengan genre musik. Hasilnya, sistem ini dapat memberikan rekomendasi yang sesuai dengan emosi seperti bahagia, sedih, atau netral. Dataset FER-2013 digunakan untuk melatih model ini, menghasilkan akurasi validasi hingga 97%. Selain itu, pendekatan berbasis data sosial seperti tag dari platform musik (contoh: Spotify atau Last.fm) juga digunakan untuk merekomendasikan musik berdasarkan suasana hati pengguna. Penelitian ini mencatat bahwa data pengguna seperti playlist, genre favorit, dan

ulasan dapat menciptakan rekomendasi yang lebih akurat. Ho et al. (2024) mengembangkan Emotion-Based Movie Recommender System (E-MRS) yang memanfaatkan emosi pengguna sebagai parameter utama. Emosi pengguna diidentifikasi melalui kombinasi warna dan preferensi pengguna sebelumnya. Sistem ini berhasil mengatasi masalah cold start dan memperbaiki relevansi rekomendasi dengan pendekatan hybrid (CBF dan CF).

Studi lain oleh Burke (2002) mengkaji pendekatan berbasis konten untuk merekomendasikan film berdasarkan aktor, sutradara, dan sinopsis. Namun, pendekatan ini memiliki keterbatasan dalam menangkap dinamika preferensi pengguna yang dipengaruhi oleh emosi.

Efektivitas rekomendasi dievaluasi menggunakan metrik seperti precision, recall, dan novelty. Bang et al. (2024) mencatat bahwa sistem rekomendasi berbasis emosi menunjukkan tingkat user satisfaction yang lebih tinggi dibandingkan sistem tradisional, terutama karena personalisasi yang lebih baik.

2.1.5 Evaluasi Sistem Rekomendasi

Berbagai metode evaluasi telah diterapkan dalam penelitian sistem rekomendasi, termasuk metrik kuantitatif seperti *Precision*, *Recall*, *F1-Score*, *Mean Absolute Error (MAE)*, dan *Root Mean Squared Error (RMSE)*. Tripathi et al. (2018) menekankan pentingnya evaluasi dengan metrik ini untuk menentukan akurasi model analisis emosi berbasis teks.

Selain itu, Bang et al. (2024) menunjukkan pentingnya mempertimbangkan metrik kualitas rekomendasi seperti *novelty*, *diversity*, dan *serendipity* untuk memastikan pengalaman pengguna yang holistik.

2.1.6 Identifikasi Gap Penelitian

Dari berbagai penelitian di atas, beberapa celah yang dapat diidentifikasi adalah:

1. Representasi emosi dari teks panjang seperti cerita atau artikel masih kurang dieksplorasi.
2. Sistem rekomendasi berbasis emosi sering kali tidak mengintegrasikan variasi metrik kualitas seperti *diversity* dan *serendipity*.

Penelitian ini berusaha mengatasi celah-celah tersebut dengan mengembangkan sistem rekomendasi berbasis emosi yang menggunakan data teks pendek dan panjang, serta mengevaluasi model menggunakan metrik akurasi dan kualitas rekomendasi.

2.2 Landasan Teori

Landasan teori pada penelitian ini mencakup konsep-konsep utama yang menjadi dasar pengembangan sistem rekomendasi berbasis emosi. Beberapa teori dan pendekatan utama bisa dilihat sebagai berikut:

2.2.1 Sistem Rekomendasi

Sistem rekomendasi adalah teknologi berbasis algoritma yang membantu pengguna menemukan item yang relevan di tengah banyaknya pilihan. Teknologi ini digunakan di berbagai industri, termasuk e-commerce, hiburan, dan pendidikan.

1. Content-Based Filtering (CBF)

CBF bekerja dengan membangun profil pengguna berdasarkan preferensi mereka terhadap atribut item tertentu. Sebagai contoh, pada platform streaming, atribut seperti genre film, aktor, atau sutradara dapat digunakan untuk merekomendasikan film serupa yang telah dinikmati pengguna sebelumnya. Keunggulan:

- Tidak bergantung pada data pengguna lain sehingga cocok untuk pengguna baru.
 - Mudah diimplementasikan jika deskripsi item tersedia.
- Kelemahan:

2. Collaborative Filtering (CF)

CF menggunakan data interaksi antar pengguna untuk memberikan rekomendasi. Misalnya, jika dua pengguna memiliki pola rating yang mirip, item yang disukai oleh salah satu pengguna kemungkinan akan direkomendasikan kepada pengguna lainnya. CF terbagi menjadi:

- *User-based*: Menganalisis kesamaan antar pengguna.
 - *Item-based*: Menganalisis kesamaan antar item.
- Kelemahan utama adalah masalah sparsity, yaitu ketika data interaksi pengguna terlalu sedikit untuk menghasilkan rekomendasi yang akurat.

3. Hybrid Filtering

Pendekatan hybrid menggabungkan CBF dan CF untuk memanfaatkan keunggulan keduanya sekaligus mengurangi kelemahannya. Contohnya adalah algoritma Singular Value Decomposition (SVD) yang sering digunakan untuk mengatasi masalah sparsity dengan mereduksi dimensi data.

2.2.2 Analisis Emosi

Analisis emosi adalah proses untuk mengidentifikasi dan memahami keadaan emosional seseorang melalui data yang dihasilkan, seperti teks, gambar, atau suara. Emosi memiliki peran penting dalam menentukan preferensi individu terhadap item tertentu, sehingga pengintegrasian analisis emosi dalam sistem rekomendasi dapat meningkatkan personalisasi.

1. Model Emosi Plutchik

Robert Plutchik (1980) mengusulkan model roda emosi (Plutchik's Wheel of Emotions) yang menggambarkan delapan emosi dasar: bahagia, sedih, takut, marah, terkejut, jijik, antisipasi, dan percaya. Emosi-emosi ini saling berhubungan dan dapat digabungkan untuk membentuk emosi kompleks. Misalnya, cinta adalah kombinasi dari bahagia dan percaya. Model ini memberikan kerangka dasar untuk menganalisis dan mengklasifikasikan emosi dalam sistem berbasis AI.

2. Analisis Emosi Berbasis Teks

Analisis teks menjadi salah satu cara paling fleksibel untuk mendeteksi emosi. Teknik seperti *TF-IDF* dan *Word Embedding* digunakan untuk mengonversi teks menjadi representasi numerik yang dapat diproses oleh algoritma pembelajaran mesin. Beberapa model seperti *BERT* dan *XLNet* telah menunjukkan hasil yang sangat baik dalam analisis emosi berbasis teks.

3. Analisis Multimodal

Pendekatan multimodal menggabungkan data dari berbagai sumber, seperti teks, gambar, dan suara, untuk mendeteksi emosi dengan lebih akurat. Misalnya, sistem rekomendasi dapat menggunakan ekspresi wajah pengguna bersama teks yang mereka tulis untuk memberikan hasil yang lebih personal.

2.2.3 Algoritma untuk Sistem Rekomendasi

1. Cosine Similarity

Cosine Similarity mengukur kesamaan antara dua vektor dengan menghitung kosinus sudut di antara mereka. Ini sering digunakan dalam sistem rekomendasi berbasis konten untuk membandingkan profil pengguna dengan fitur item:

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}},$$

Keunggulan: Tidak bergantung pada panjang vektor, sehingga cocok untuk data teks seperti deskripsi film atau ulasan.

2. Euclidean Distance

Euclidean Distance menghitung jarak geometris antara dua titik dalam ruang multidimensi:

$$\text{Euclidean Distance} = |X - Y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Algoritma ini berguna untuk sistem rekomendasi berbasis metrik jarak, di mana item dengan jarak terpendek dianggap paling relevan.

3. K-Nearest Neighbors (KNN)

KNN mencari k item terdekat dengan menggunakan metrik seperti Euclidean Distance atau Cosine Similarity. Dalam konteks rekomendasi, KNN digunakan untuk mencari item yang paling mirip dengan preferensi pengguna. Keunggulannya adalah fleksibilitas dalam menangani data kompleks, namun memiliki kelemahan dalam skala besar karena biaya komputasi yang tinggi.

2.2.4 Evaluasi Sistem Rekomendasi

Evaluasi adalah langkah penting untuk menentukan efektivitas dan kualitas sistem rekomendasi. Beberapa metrik yang digunakan meliputi:

1. Precision

Precision mengukur proporsi rekomendasi yang relevan terhadap semua rekomendasi yang diberikan:

$$Precision = \frac{TP}{TP + FP}$$

2. Recall

Recall mengukur sejauh mana sistem dapat merekomendasikan item yang relevan:

$$Recall = \frac{TP}{TP + FN}$$

3. Mean Absolute Error (MAE)

MAE menghitung rata-rata kesalahan absolut antara nilai prediksi dan nilai aktual:

$$mae = \frac{\sum_{i=1}^n abs(y_i - \lambda(x_i))}{n}$$

4. Diversity dan Novelty

Metrik ini mengevaluasi variasi dan kebaruan rekomendasi, yang penting untuk meningkatkan kepuasan pengguna dalam jangka panjang. Misalnya, rekomendasi dengan *diversity* tinggi mengurangi monoton dalam pengalaman pengguna.

BAB 3 DESAIN DAN IMPLEMENTASI SISTEM

3.1 Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dan kualitatif. Pendekatan kuantitatif digunakan untuk mengevaluasi performa sistem rekomendasi berbasis emosi dengan metrik seperti *Precision*, *Recall*, *F1-Score*, *Mean Absolute Error (MAE)*, dan *Root Mean Squared Error (RMSE)*. Pendekatan kualitatif digunakan untuk memahami pengalaman pengguna melalui survei kepuasan pengguna (*User Satisfaction*) dan Click-Through Rate (CTR).

3.2 Tahapan Penelitian

Tahapan penelitian dirancang untuk memastikan proses pengembangan sistem rekomendasi berjalan sistematis, terdiri dari lima tahap utama:

3.2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini berasal dari dua sumber utama:

1. **Data Teks Pendek (Short Text):** Dataset berupa tweet dengan hashtag emosi seperti #happy, #sad, atau #angry, yang digunakan sebagai label supervisi.
2. **Data Teks Panjang (Long Text):** Dataset berupa headline berita atau cerita anak-anak yang dianotasi berdasarkan enam emosi dasar (bahagia, sedih, marah, takut, jijik, terkejut).

3.2.2 Preprocessing Data

Data yang telah dikumpulkan akan diproses untuk meningkatkan kualitas analisis. Proses ini meliputi:

- **Pembersihan Data:** Menghapus simbol, tanda baca, atau kata yang tidak relevan.
- **Tokenisasi:** Membagi teks menjadi unit kata atau frasa.
- **Representasi Data:** Menggunakan metode *TF-IDF* untuk merepresentasikan data dalam bentuk vektor numerik.

3.2.3 Pengembangan Model

Model rekomendasi dikembangkan menggunakan pendekatan Content-Based Filtering (CBF) dengan tiga algoritma utama:

1. **Cosine Similarity:** Digunakan untuk mengukur kesamaan antara profil pengguna dan atribut item.
2. **Euclidean Distance:** Digunakan untuk mengukur jarak antara item dalam ruang fitur.
3. **K-Nearest Neighbors (KNN):** Digunakan untuk menemukan item paling relevan berdasarkan metrik kesamaan.

3.2.4 Implementasi Sistem

Sistem rekomendasi akan diimplementasikan dalam dua tahap:

1. **Pengembangan Prototype:** Sistem akan dibangun menggunakan Python dengan pustaka seperti Scikit-learn untuk algoritma dan Pandas untuk pengolahan data.
2. **Uji Coba Sistem:** Prototype akan diuji menggunakan dataset yang telah diproses untuk mengukur performa model.

3.2.5 Evaluasi Sistem

Evaluasi dilakukan untuk menilai efektivitas sistem rekomendasi berdasarkan:

1. **Metrik Kuantitatif:** Menggunakan metrik evaluasi seperti Precision, Recall, F1-Score, MAE, dan RMSE untuk mengukur akurasi dan efisiensi algoritma.
2. **Metrik Kualitatif:** Menggunakan survei pengguna untuk menilai kepuasan, Click-Through Rate (CTR), dan Conversion Rate.

3.3 Alat dan Bahan

- **Bahasa Pemrograman:** Python.
- **Pustaka Python:** Scikit-learn, Numpy, Pandas, Matplotlib.
- **Platform Pengujian:** Google Colab atau Jupyter Notebook.
- **Dataset:** Data emosi dari tweet dan headline berita.

3.4 Desain Eksperimen

Eksperimen dilakukan untuk menguji sistem rekomendasi berbasis emosi dalam berbagai skenario:

1. **Skenario 1:** Menganalisis pengaruh algoritma Cosine Similarity terhadap akurasi rekomendasi.

2. **Skenario 2:** Membandingkan performa Euclidean Distance dan KNN dalam mendeteksi item relevan.
3. **Skenario 3:** Mengukur dampak keberagaman genre terhadap kepuasan pengguna.

3.5 Rancangan Evaluasi

Evaluasi dilakukan dalam tiga tahap:

1. **Evaluasi Akurasi:** Menggunakan metrik kuantitatif untuk mengukur performa algoritma.
2. **Evaluasi Kualitas Rekomendasi:** Menggunakan metrik seperti novelty, diversity, dan serendipity untuk menilai kualitas rekomendasi.
3. **Evaluasi Kepuasan Pengguna:** Melalui survei dan pengamatan terhadap Click-Through Rate (CTR) dan Conversion Rate.

3.6 Desain Sistem

Desain sistem rekomendasi berbasis emosi ini bertujuan untuk menyediakan rekomendasi personal yang relevan berdasarkan emosi pengguna. Sistem dirancang dengan pendekatan modular untuk memudahkan pengembangan, pengujian, dan pemeliharaan.

3.6.1 Arsitektur Sistem

Sistem terdiri dari beberapa komponen utama, seperti yang disusun berikut ini:

1. **Antarmuka Pengguna (Frontend)**

Antarmuka ini memungkinkan pengguna untuk memasukkan teks, seperti ulasan atau opini. Sistem juga menyediakan antarmuka untuk menampilkan hasil rekomendasi berupa genre musik dan film.

2. **Modul Preprocessing Data**

Modul ini membersihkan data teks yang dimasukkan pengguna. Langkah-langkah yang dilakukan:

- Menghapus simbol dan tanda baca yang tidak relevan.
- Menggunakan *stemming* dan *lemmatization* untuk mengurangi kata ke bentuk dasar.

3. **Modul Deteksi Emosi**

Modul ini menggunakan teknik pemrosesan bahasa alami (NLP) untuk mendeteksi emosi pengguna berdasarkan teks. Prosesnya meliputi:

- Representasi data teks menggunakan *TF-IDF*.
- Penerapan model pembelajaran mesin (SVM atau CRF) untuk klasifikasi emosi.

4. Modul Sistem Rekomendasi

Sistem rekomendasi memanfaatkan Content-Based Filtering dengan tiga algoritma utama:

- **Cosine Similarity:** Membandingkan profil pengguna dengan atribut item.
- **Euclidean Distance:** Mengukur jarak antara vektor fitur item dan pengguna.
- **K-Nearest Neighbors (KNN):** Menemukan item paling relevan berdasarkan jarak terdekat. Modul ini akan mengakses basis data musik dan film untuk menghasilkan rekomendasi.

5. Modul Evaluasi Sistem

Modul ini mengukur performa sistem dengan metrik evaluasi seperti Precision, Recall, F1-Score, dan MAE.

6. Basis Data

Basis data menyimpan metadata musik dan film, termasuk deskripsi, genre, dan atribut lainnya, serta log interaksi pengguna.

3.6.2 Diagram Alur Sistem

Alur sistem dirancang untuk memastikan setiap komponen berjalan sesuai fungsi. Diagram flowchat dapat disusun sebagai berikut:

1. Pengguna memasukkan teks

- Input teks diproses oleh modul preprocessing untuk pembersihan dan tokenisasi.

2. Analisis emosi dilakukan

- Teks yang telah diproses diklasifikasikan menggunakan model NLP untuk mendeteksi emosi utama (misalnya: bahagia, sedih, marah).

3. Pemrosesan rekomendasi

- Profil pengguna dibuat berdasarkan emosi yang terdeteksi.

- Algoritma Content-Based Filtering digunakan untuk mencari item yang sesuai dengan profil pengguna.

4. Hasil rekomendasi ditampilkan

- Sistem menampilkan daftar rekomendasi musik dan film yang relevan melalui antarmuka pengguna.

3.6.3 Representasi Data dan Penyimpanan

Sistem menggunakan representasi data yang terstruktur untuk memudahkan pemrosesan:

- **Metadata Musik dan Film:**

Tabel atribut mencakup ID, nama, genre, dan deskripsi.

- **Data Pengguna:**

Tabel interaksi menyimpan log emosi dan item yang direkomendasikan.

3.6.4 Teknologi yang Digunakan

1. Bahasa Pemrograman:

- Python untuk backend dan pemrosesan data.

2. Framework:

- Flask/Django untuk pengembangan aplikasi web.

3. Pustaka NLP:

- Scikit-learn, NLTK, dan SpaCy.

4. Basis Data:

- MySQL atau PostgreSQL untuk penyimpanan data.

3.6.5 Use Case Diagram

Skenario utama yang dicakup oleh sistem adalah:

1. Pengguna Menginput Teks:

- Pengguna memasukkan teks melalui antarmuka pengguna.

2. Sistem Mendeteksi Emosi:

- Sistem memproses teks dan mengidentifikasi emosi utama.

3. Sistem Memberikan Rekomendasi:

- Sistem menghasilkan daftar rekomendasi musik dan film berdasarkan emosi.

4. Pengguna Memberikan Umpan Balik:

- Pengguna dapat memberi rating terhadap hasil rekomendasi untuk meningkatkan performa sistem.