**Laporan Tugas Besar Sistem Pemberi Rekomendasi**

Diajukan untuk memenuhi Tugas Mata Kuliah Sistem Pemberi Rekomendasi

Dosen Pengampu: Agung Toto Wibowo, S.T., M.T.



Disusun oleh :

|  |  |
| --- | --- |
| Muhammad Rizki Nurfiqri | (1301204009) |
| Satria Ginting | (1301204038) |

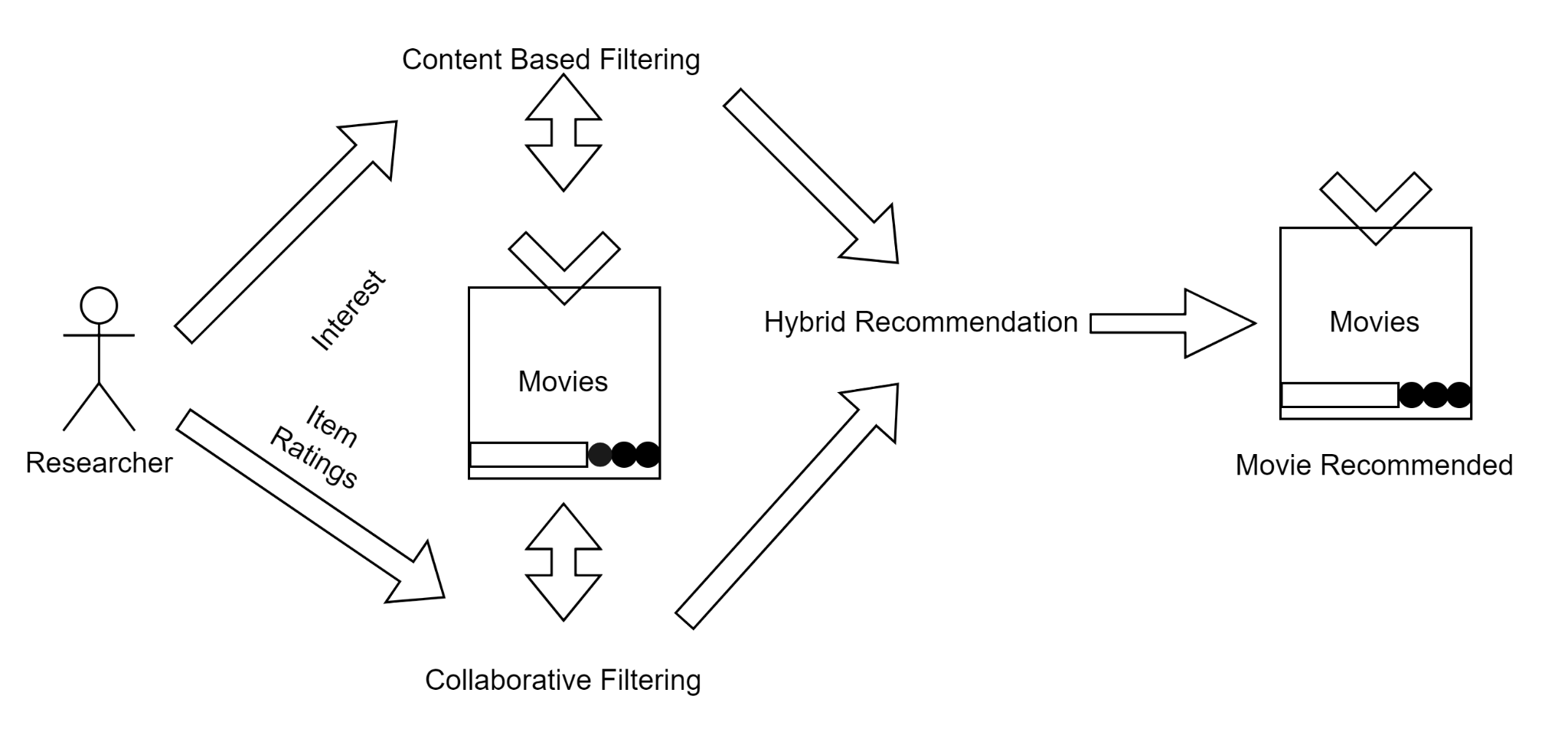
**PROGRAM STUDI INFORMATIKA**

**FAKULTAS INFORMATIKA**

**TELKOM UNIVERSITY**

**2022**

# Hybrid Recommend System



*Hybrid recommender systems* digunakan untuk meningkatkan kualitas model dan membuat rekomendasi yang lebih efektif.

*Content Based Filtering*, memberikan rekomendasi berdasarkan kemiripan atribut dari item atau barang yang disukai.

*Collaborative Filtering*, memberikan rekomendasi berdasarkan feedback dari user yang lain atau dari diri sendiri.

# Penjelasan mengenai algoritma yang dilakukan

Formulasi Masalah

Desain sebuah mekanisme rekomendasi *Hybrid* yang melibatkan *Collaborative Filtering*, *Content Based Filtering* pada dataset Food.com – Recipes and Reviews. Ini berisi 1257 resep dan 4010 Review.

Eksplorasi dan Persiapan Data

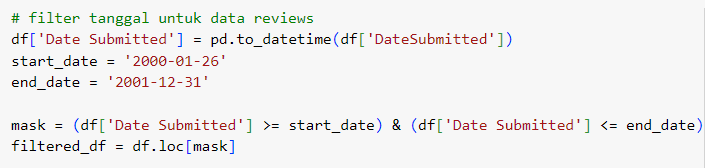
Berikut adalah proses yang terlibat untuk melakukan eksplorasi dan persiapan untuk data:

* *Data Exploration*

Seperti melakukan input library, pemanggilan dataset, mengecek info dataset, menghapus kolom yang tidak dibutuhkan. Menyaring tanggal untuk data reviews dan menyaring kolom yang saling berhubungan pada data resep.



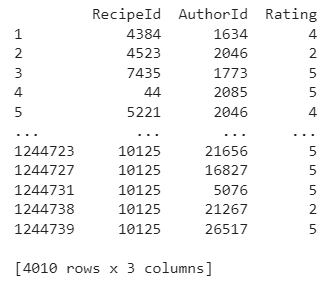
 Memanggil atau membaca dataset

menyaring dan menampilkan kolom-kolom yang sudah di saring / filter dan

A screenshot of a computer

Description automatically generatedmenyaring kolom kolom yang saling berhubungan

menampilkan beberapa baris pertama dari dataframe akhir



Filter interaksi untuk hanya menyertakan interaksi yang terkait dengan resep yang difilter dan Filter lebih lanjut interaksi untuk hanya menyertakan interaksi dengan AuthorId yang cocok lalu menampilkan hasilnya.

* *Pemodelan (Hybrid Recommender System)*
* *Most Popular*

 pendekatan sederhana di mana item yang paling populer atau paling sering digunakan oleh sebagian besar pengguna direkomendasikan kepada pengguna lain yang memiliki profil serupa. Konsep ini didasarkan pada asumsi bahwa item yang populer atau sering digunakan oleh banyak orang cenderung disukai oleh pengguna lain juga.

**

* *Collaborative Filtering*

Memberikan rekomendasi berdasarkan feedback dari user yang lain atau dari diri sendiri. Penerapan dalam rekomendasi resep dan reviews yaitu berdasarkan perilaku dan preferensi pengguna lain yang serupa.

* + SVD (Singular Value Decomposition)

memecah matriks rating menjadi matriks yang lebih sederhana untuk memprediksi rating yang mungkin diberikan oleh pengguna terhadap item yang belum dinilai.

A screenshot of a computer

Description automatically generated

* + *KNNBasic*

Algoritma KNNBasic menggunakan pendekatan neighbor-based untuk memberikan rekomendasi. Ini mencari kesamaan antara pengguna atau item berdasarkan matriks rating dan melakukan prediksi berdasarkan rating dari pengguna terdekat.

A screenshot of a computer

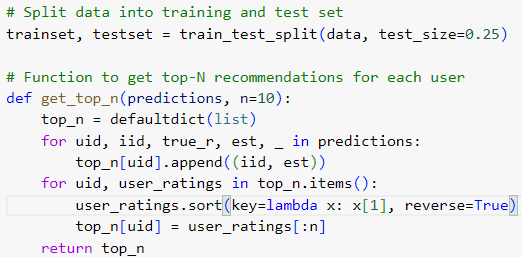
Description automatically generated

* + *Matrix Factorization with Baseline (BaselineOnly)*

Algoritma BaselineOnly menggunakan pendekatan sederhana berdasarkan nilai rata-rata atau baseline. Ini menghitung rata-rata dari semua rating, rata-rata rating per pengguna, dan rata-rata rating per item sebagai "baseline" dan menggunakan informasi ini untuk membuat prediksi.

A screenshot of a computer

Description automatically generated

Setiap algoritma dilatih dengan data pelatihan (trainset) dan diuji dengan data pengujian (testset). Kemudian, menggunakan fungsi get\_top\_n, setiap algoritma menghasilkan 10 rekomendasi teratas untuk setiap pengguna berdasarkan estimasi rating yang diberikan oleh algoritma untuk item yang belum dinilai oleh pengguna. 

A computer screen shot of a computer code

Description automatically generatedSetelah itu, hasil rekomendasi teratas dari masing-masing algoritma disimpan dalam top\_n\_recommendations dalam bentuk DataFrame yang berisi AuthorId, RecipeId, dan EstimatedRating untuk setiap rekomendasi. Akhirnya, hasil rekomendasi teratas dari setiap algoritma ditampilkan dalam bentuk tabel yang menunjukkan 10 rekomendasi teratas untuk setiap algoritma.

* Selanjutnya penggabungan algoritma SVD, KNNBasic, dan BaselineOnly.
* Data dimuat menggunakan Dataset.load\_from\_df dari final\_df, yang berisi kolom 'RecipeId', 'AuthorId', dan 'Rating'.
* Data dibagi menjadi set pelatihan (trainset) dan set pengujian (testset) menggunakan train\_test\_split.



A screenshot of a computer program

Description automatically generated

* Algoritma-algoritma (SVD, KNNBasic, BaselineOnly) dijalankan dengan data pelatihan dan diuji dengan set pengujian.
* Performa tiap algoritma diukur menggunakan metrik Root Mean Squared Error (RMSE) dan Mean Absolute Error (MAE) dengan accuracy.rmse() dan accuracy.mae().
* Semua hasil rekomendasi dari berbagai algoritma digabungkan menjadi satu DataFrame (combined\_recommendations).
* Data yang telah diambil rata-ratanya diurutkan untuk mendapatkan rekomendasi teratas secara keseluruhan dari semua algoritma.
* 10 rekomendasi teratas secara keseluruhan disimpan dalam overall\_top\_10\_df.
* Rata-rata RMSE dan MAE dari semua algoritma dihitung untuk mengetahui kinerja rata-rata dari semua algoritma yang digunakan.

A screenshot of a computer

Description automatically generated

* *Content Based Filtering*

Memberikan rekomendasi berdasarkan kemiripan atribut dari item atau barang yang disukai. Pada sistem rekomendasi food.com untuk menganalisis kesamaan antar resep pada sebuah DataFrame dalam menentukan seberapa mirip atau relevan satu resep dengan resep lainnya berdasarkan konten teks seperti deskripsi, kata kunci, dan bahan-bahan yang digunakan dalam resep, menggunakan teknik TF-IDF dan cosine similarity.

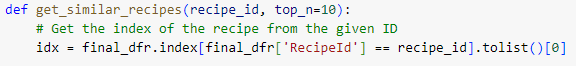
A screen shot of a computer program

Description automatically generatedmengubah teks menjadi vektor TF-IDF, mengubah teks yang telah digabungkan menjadi representasi vektor TF-IDF. menghitung skor kesamaan cosine antara setiap pasangan resep berdasarkan vektor TF-IDF mereka.

Mencetak DataFrame cosine\_sim\_df, yang berisi skor kesamaan cosine antar resep berdasarkan analisis TF-IDF



get\_similar\_recipes yang bertujuan untuk mendapatkan rekomendasi resep yang mirip berdasarkan skor kesamaan kosinus yang telah dihitung sebelumnya menggunakan TF-IDF.



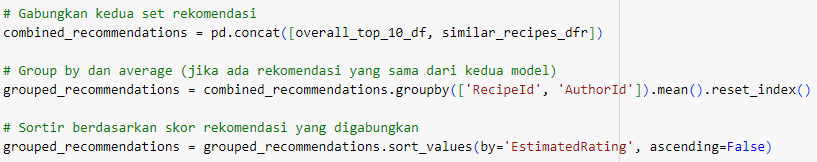
Menampilkan hasil resep-resep teratas yang mirip dengan resep yang memiliki RecipeId 487

A close-up of a white background

Description automatically generated

* Hybrid Recommendation

menggabungkan rekomendasi dari dua model berbeda, yaitu collaborative filtering dan content-based filtering, untuk menghasilkan rekomendasi "hibrida" yang merupakan kombinasi dari kedua jenis rekomendasi tersebut.



Menampilkan 10 rekomendasi teratas Menggunakan fungsi top\_10\_hybrid\_recommendations

A screenshot of a computer code

Description automatically generated

Hasil dari 10 Hybrid Recommendations

A screenshot of a computer

Description automatically generated