Rekomendasi K-Drama Dengan Menggunakan

Metode MinMaxScaler, NearestNeighbors,

dan cosine\_similarity

Khurin Rahmadini  
Fakultas Sains dan Teknologi  
Program Studi Teknik Informatika   
Universitas Darussalam Gontor Kampus Putri   
Sambirejo, Mantingan, Ngawi, Jawa Timur 63260   
Khurinrahmadini@mhs.unida.gontor.ac.id

*Abstract*—Perekomendasian K-Drama dengan menggunakan metode MinMaxScaler, NearestNeightBors, dan Cosine Similarity dapat memberikan rekomendasi sesuai informasi yang diberikan user namun tidak dapat memberikan rekomendasi diluar dari dataset yang telah ada.

Keywords—MinMaxScaler, NearestNeighbors, Cosine Similarity

# Pendahuluan

Sistem rekomendasi telah banyak digunakan dalam *e-commerce* untuk merekomendasikan barang-barang seperti buku, musik dan artikel. Sebagai salah satu contohnya adalah sistem rekomendasi yakni rekomendasi film yang mampu memberikan *review* kepada pengguna. Pada umumnya rekomendasi yang dilakukan berdasarkan *review* untuk film ataupun produk lain di media *online* atau *digital library* hanya dilakukan dengan cara memilih dari beberapa kata yang merepresentasikan *review* dari setiap pengguna[1].

Sistem rekomendasi saat ini telah menjadi bagian penting dari industri, karena dapat digunakan untuk membantu menciptakan strategi promosi penjualan di banyak situs online dan aplikasi selular. Sebagai contoh 60 persen video klik di Youtube berasal dari rekomendasi . saat ini banyak perusahaan menggunakan metode deep learning untuk meningkatkan kualitas rekomendasi. Keberhasilan metode deep learning untuk rekomendasi, baik dalam sains dan industri, memerlukan tinjauan komprehensif dan kesimpulan yang kuat[3].

Secara psikologis, manusia suka dengan drama dan senang kejutan, makanya permainan roller coaster ada dan antreannya mengular. Campur aduk emosi inilah yang diberikan oleh drama korea. Sebuah pengalaman emosi yang membuat kamu memamah manis-pahit-asam-kecut-getir dunia percintaan, lengkap dengan tingkah-polah karakter yang unyu dan penampilan yang rupawan. Namun ada pula beberapa alasan yang membuat penonton K-Drama semakin banyak, yaitu Daya Tarik Budaya, mempelajari hal baru, percintaan yang sweet, fashionable, sajian makanan yang menggugah selera, mudah diakses, dan asyik didiskusikan bersama teman[2]. Oleh karena banyaknya penggemar K-Drama pada paper kali ini saya akan membahas paper dengan mengusung judul mengenai Rekomendasi K-Drama Dengan Menggunakan Metode MinMaxScaler, Nearest Neighbors,dan Cosine Similarity.

Dalam pembahasan paper ini digunakan beberapa metode diantara, yaitu metode MinMaxScaler, Nearest Neighbors, dan Cosine Similarity. Dimana metode MinMaxScaler berfungsi untuk melakukan transformasi data linear terhadap data asli.

# Metode

Dalam pembuatan penelitian ini penulis menggunakan metode *MinMaxScaler, NearestNeighbors, dan Cosine Similarity* yang mana metode ini merupakan salah satu metode dalam deep learning [3].

## MinMaxScaler

Min-Max Scaling, yang sering dikenal juga dengan normalisasi data atau normalization. Min-Max Scaling bekerja dengan scaling data atau menyesuaikan data dalam rentang atau range tertentu (range nilai minimum hingga nilai maksimum), dengan rentang yang biasa digunakan adalah 0 hingga 1.

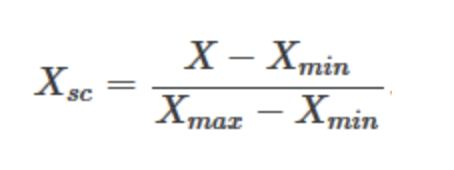


Fig.1. Rumus MinMaxScaler

## Nearest Neighbors

Merupakan algoritma yang menggunakan teknik statistika untuk menemukan sekumpulan pengguna, dikenal sebagai tetangga (neighbor) yang memiliki pendapat sama atau setuju terhadap pengguna yang menjadi sasaran. Setelah sekumpulan tetangga terbentuk, sistem menggunakan algoritma yang berbeda untuk menggabungkan kesukaan neighbours untuk menghasilkan suatu prediksi atau rekomendasi N-teratas untuk pengguna aktif [4].

Dalam pengenalan pola, algoritma KNN merupakan metode non parametrik yang digunakan untuk klasifikasi dan regresi [5]. Algoritma KNN memiliki kelebihan yaitu dapat menghasilkan data yang kuat atau jelas dan efektif jika digunakan pada data yang besar.

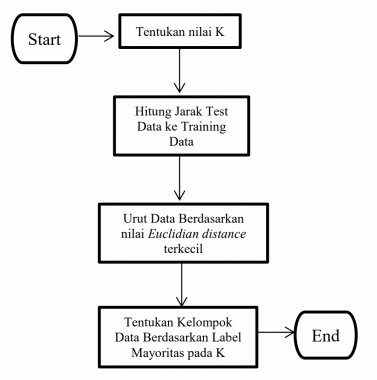


Fig.2. Flowchart KKN

Penentuan nilai k yang baik bergantung pada data yang digunakan. Umumnya, nilai k yang lebih besar mengurangi efek *noise* pada klasifikasi, namun menyebabkan batasan antar kelas sedikit berbeda. Jarak antar data yang digunakan dalam algoritma KNN dapat dihitung menggunakan rumus jarak *Euclidean Distance* dan *Manahattan Distance*. Rumus jarak yang biasa digunakan untuk KNN adalah *Euclidean Distance.*

K-Nearest Neighbor (KNN) adalah suatu metode yang menggunakan algoritma supervised dimana hasil dari query instance yang baru diklasifikan berdasarkan mayoritas dari kategori pada KNN. Algoritma metode KNN sangatlah sederhana, bekerja dengan berdasarkan pada jarak terpendek dari sample uji ke sample latih untuk menentukan KNN nya. Setelah mengumpulkan KNN, kemudian diambil mayoritas dari KNN untuk dijadikan prediksi dari sample uji [6]

## C. Cosine similarity

Metode cosine similarity merupakan metode yang digunakan untuk menghitung similarity (tingkat kesamaan) antar dua buah objek. Secara umum perhitungan metode ini

didasarkan pada vector space similarity measure. Metode cosine similarity ini menghitung similarity antara dua buah objek (misalkan D1 dan D2) yang dinyatakan dalam dua buah vektor dengan menggunakan kata kunci dari sebuah dokumen sebagai ukuran [7]. Persamaan 1 di bawah ini adalah rumus perhitungan cosine similarity.

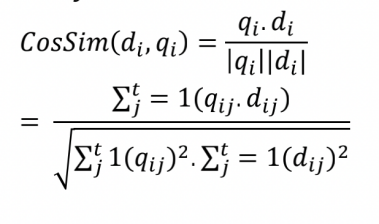


Fig.3. Perhitungan Cosine Similarity

Pengukuran cosine similarity merupakan metode yang digunakan untuk menghitung tingkat kesamaan antar dua buah objek berdasarkan besar sudut cosinus. Nilai sudut cosinus antara dua vektor menentukan kesamaan dua buah objek yang dibandingkan dimana nilai terkecil adalah 0 dan nilai terbesar adalah 1

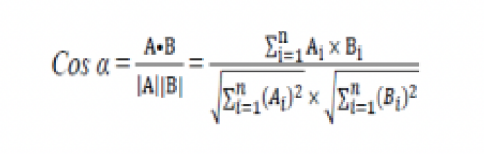


Fig.4. Cosine Similarity

*Cosine Similarity* adalah salah satu algoritma di dalam teks mining, yang berfungsi untuk mengklasifikasikan suatu dokumen atau teks. *Cosine similarity* mempunyai konsep normalisasi panjang *vektor* dengan membandingkan dua

dokumen antara dokumen A dan B

# Hasil

Hasil dari Penelitian tersebut adalah sebagai berikut:

## Import Drive

Direktori yang digunakan dalam membuat project kali ini adalah *Google Drive*, karena pembuatannya yang *Online* membuat *Drive* butuh di*import* kedalam *Google Colab.* Seluruh *dataset* yang akan digunakan dalam *project* ini disimpan dalam suatu folder di *Google Drive*, baik berupa Citra atau Tabel.

## Import Library

Pada project ini dibutuhkan 5 buah library yaitu library *numpy, pandas, seaborn, matplotlib, dan scikit learn*. .

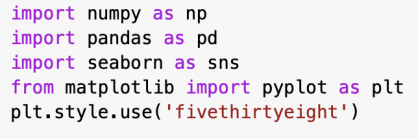


Fig.3. library yang dibutuhkan

## Memanggil Model Yang Ada Di Scikit Learn

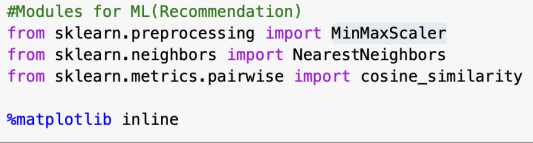


Fig.4. Syntax Untuk Meanggil Scikit Learn

## Menentukan Fitur Yang Ada Pada Rekomendasi

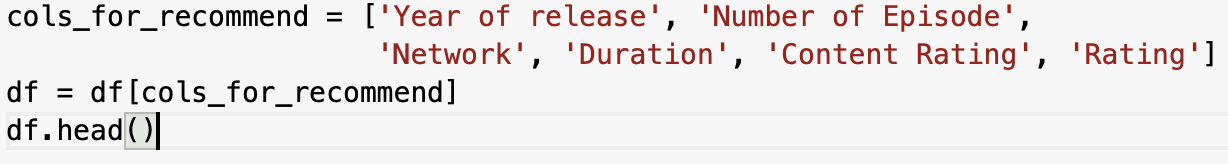


Fig.5. Fitur Rekomendasi

## Mencari Rekomendasi Drama

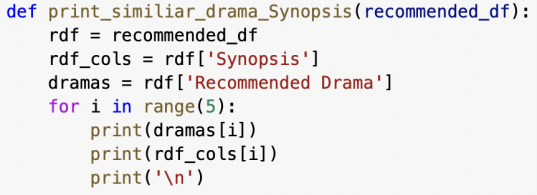


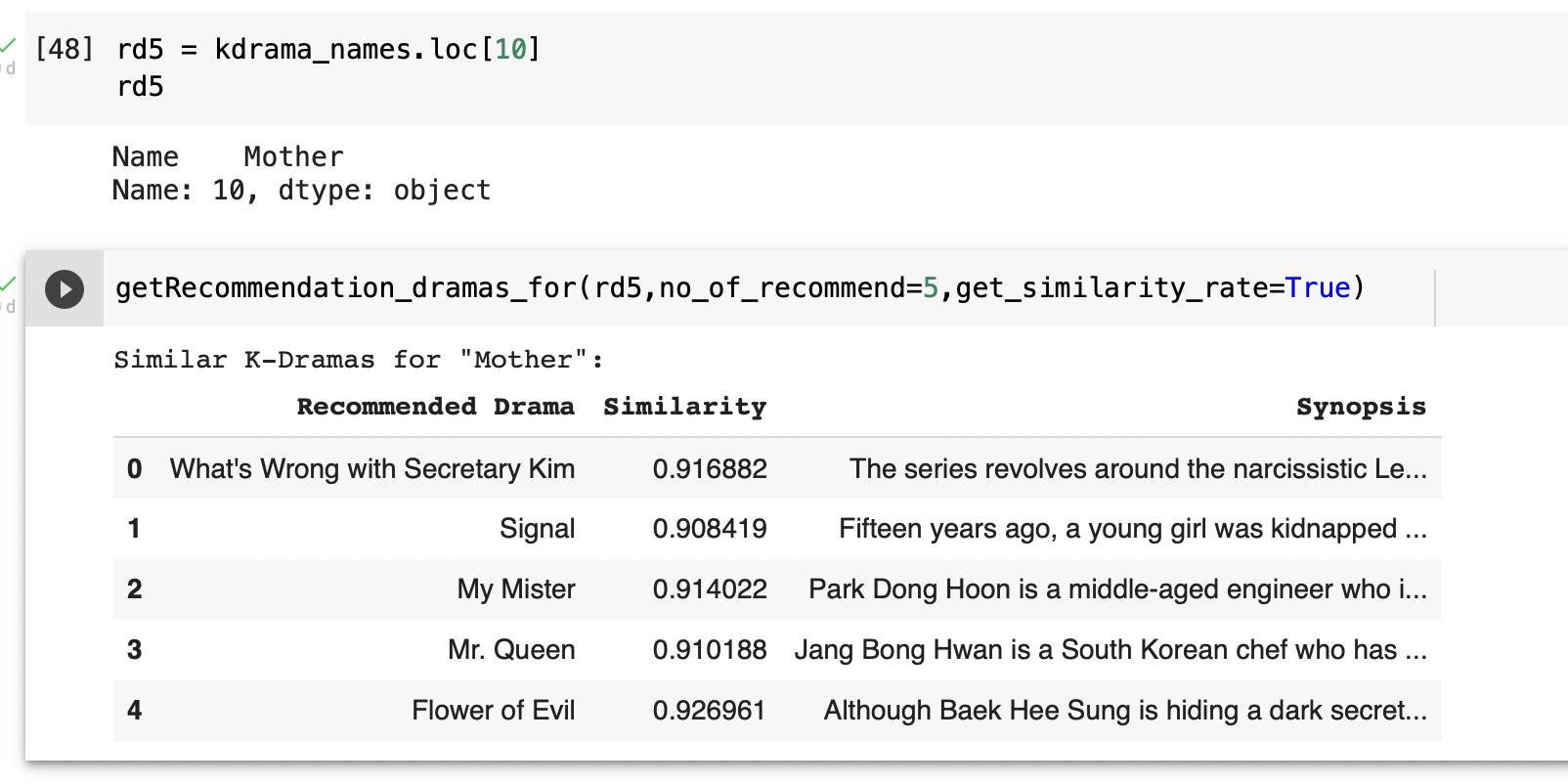
Fig.6. Syntax Rekomendasi Drama

## G. Ringkasan

Keseluruhan dari tahap-tahap diatas merupakan keseluruhan cara-cara untuk menentukan Rekomendasi film berdasarkan dataset yang telah di buat.

## H. Hasil

Hasil dari percobaan tersebut sebagai berikut:



# Kesimpulan

Hasil dari penerapan ketiga metode yang digunakan melakukan pemrosesan rekomendasi berdasarkan apa yang di minta oleh user dengan mencari data pada dataset yang telah disediakan. Kelemahan dari penelitian ini yaitu program tidak dapat memberikan rekomendasi diluar dari dataset yang telah diatur

##### References

1. Ningrum. L, “Pendekatan Deep Learning Untuk Sistem Rekomendasi Review Film”. Universitas Gunadarma
2. Dr. Rizal Fadli, “Alasan Psikologi Mengapa Drama Korea Begitu Digemari”. Diakses pada 17.50 pada Senin, 10 Oktober 2021
3. Ikasari. Diana, Suhendra. Adang, Farida. Nuke, “Metode Deep Learning Pada Sistem rekomendasi”, Seminar Teknologi Informasi Dan Komunikasi STI&K, Volume.2, 26 Juli 2018
4. Schafer, B., Joseph, Konstan and John Riedl, “Recommender Systems in E-Commerce,” in Minneapolis, University of Minnesota, 2007.
5. N.S. Altman, *An introduction to kernel and nearest-neighbor nonparametric regression*.: The American Statistician, 1992.
6. Samuel. Rut, Natan. Ripa, Fitria, Syafiqoh. Ummi, “Penerapan Cosine Similarity dan K-Nearest Neighbor (KKN) Pada Klasifikasi Dan Pencarian Buku”. Journal of Big Data Analytic and Artificial Intelligence, Vol. 1, No. 1, Tahun 2018
7. Pradnyana, Gede Aditra. "Perancangan dan Implementasi Automated Document Integration dengan Menggunakan Algoritma Complete Linkage Agglomerative Hierarchical Clustering." Jurnal Ilmu Komputer 5.2 (2012).

Fig.7. Hasil Akhir Dari Pencarian Rekomendasi K-Drama