

MOVIEMU : SISTEM REKOMENDASI FILM MENGGUNAKAN ALGORITMA COSINE SIMILARITY

Rizal Harjo Utomo^{1*}, I Gede Susrama Mas Diyasa²

*Email : 21083010101@student.upnjatim.ac.id

Universitas Pembangunan Nasional “Veteran” Jawa Timur^{1,2}

Studi Independen adalah inisiatif pemerintah yang diluncurkan pada tahun 2020 oleh Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek) untuk memberikan mahasiswa kesempatan untuk memperdalam pengetahuan dan keterampilan di luar lingkungan perguruan tinggi melalui berbagai program inovatif. Salah satu program unggulan dalam MBKM adalah Magang dan Studi Independen Bersertifikat (MSIB), yang memungkinkan mahasiswa untuk terlibat dalam proyek nyata di industri dan mendapatkan sertifikat sebagai pengakuan atas pencapaian mereka. Penulis mengikuti program Studi Independen Bersertifikat (MSIB) di Bangkit Academy dengan *learning path Machine Learning*, Bangkit adalah mitra dalam program MSIB, yang bekerja sama dengan Google, didukung oleh perusahaan seperti GoTo, Traveloka, dan Kemendikbudristek RI. Dengan dukungan Kampus Merdeka, Bangkit menawarkan 9.000 tempat pada tahun 2024 bagi mahasiswa Indonesia untuk memastikan kecakapan yang sesuai dengan kebutuhan industri. Program ini dipandu oleh para ahli dari perusahaan teknologi dan startup Indonesia, dan peserta akan memperoleh keahlian dalam Machine Learning, Mobile Development, atau Cloud Computing bersama sertifikasi global dari Google. Untuk memperluas pengetahuan dan pengalaman dalam bidang pemrograman Machine Learning, penulis memutuskan untuk melakukan studi independen dengan mengikuti kelas machine learning dalam program Bangkit Academy dan menyelesaikan proyek dengan judul “MOVIEMU”. Di era digital saat ini, jumlah konten film yang tersedia untuk dinikmati konsumen meningkat secara eksponensial. Hal ini menciptakan tantangan bagi pengguna untuk menemukan film yang sesuai dengan preferensi mereka. Salah satu solusi yang dapat digunakan untuk mengatasi tantangan tersebut adalah dengan mengembangkan sistem rekomendasi film yang dapat membantu pengguna menemukan film yang sesuai dengan selera mereka. “MOVIEMU” hadir sebagai solusi permasalahan tersebut dengan mengimplementasikan algoritma cosine kemiripan untuk memberikan rekomendasi film yang akurat dan relevan.

Kata kunci: MBKM Program, Bangkit Academy, Studi Independen, Machine Learning, , Sistem Rekomendasi, Film.

Abstract

Independent Study is a government initiative launched in 2020 by the Ministry of Education, Culture, Research and Technology (Kemendikbudristek) to provide opportunities for students to deepen their knowledge and skills outside the university environment through various innovative programs. One of the flagship programs in MBKM is Certified Independent Study and Internship (MSIB), which allows students to be involved in real projects in industry and earn certificates in recognition of their achievements. The author took part in the Certified Independent Study (MSIB) program at Bangkit Academy with the Machine Learning learning path. Bangkit is a partner in the MSIB program, which collaborates with Google, supported by companies such as GoTo, Traveloka, and the Indonesian Ministry of Education and Culture. With the support of Kampus Merdeka, Bangkit is offering 9,000 places by 2024 for Indonesian students to ensure skills that match industry needs. This program is guided by experts from Indonesian technology companies and startups, and participants will gain expertise in Machine Learning, Mobile Development, or Cloud Computing along with global certification from Google. To expand knowledge and experience in the field of Machine

Learning programming, the author decided to conduct an independent study by taking a machine learning class in the Bangkit Academy program and completing a project entitled "MOVIEMU". In today's digital era, the amount of film content available for consumers to enjoy has increased exponentially. This creates a challenge for users to find movies that match their preferences. One solution that can be used to overcome this challenge is to develop a film recommendation system that can help users find films that suit their tastes. "MOVIEMU" is here as a solution to this problem by implementing the cosine equation algorithm to provide accurate and relevant film recommendations.

Keywords: MBKM Program, Startup Campus, Independent Study, Machine Learning, Recommendation System, Movie.

Pendahuluan

Bapak Nadiem Anwar Makarim, sebagai Menteri Pendidikan dan Kebudayaan (Mendikbud), memperkenalkan program Kampus Merdeka, yang membuka peluang bagi mahasiswa dari berbagai perguruan tinggi, baik negeri maupun swasta. Salah satu program unggulan dalam Kampus Merdeka adalah Magang dan Studi Independen Bersertifikat (MSIB), yang bertujuan memberikan pengalaman dunia kerja kepada mahasiswa melalui kontribusi pada proyek nyata yang dibimbing oleh tim ahli.

MSIB mencakup berbagai program, termasuk yang disebut MSIB, yang memberikan peluang bagi mahasiswa untuk belajar dan mengembangkan diri di luar lingkungan kelas, tetapi tetap diakui sebagai bagian dari kurikulum. Studi independen ini memberikan keuntungan berupa pembelajaran yang relevan, pengetahuan praktis, sertifikasi sesuai kebutuhan industri, peluang karir, dan membangun jaringan profesional.

Bangkit adalah mitra dalam program MSIB, yang bekerja sama dengan Google, didukung oleh perusahaan seperti GoTo, Traveloka, dan Kemendikbudristek RI. Dengan dukungan Kampus Merdeka, Bangkit menawarkan 9.000 tempat pada tahun 2024 bagi mahasiswa Indonesia untuk memastikan kecakapan yang sesuai dengan kebutuhan industri. Program ini dipandu oleh para ahli dari perusahaan teknologi dan startup Indonesia, dan peserta akan memperoleh keahlian dalam Machine Learning, Mobile Development, atau Cloud Computing bersama sertifikasi global dari Google.

Bangkit Academy, sebagai bagian dari MSIB, menggunakan bahasa pemrograman utama Python, yang merupakan bahasa pemrograman terpopuler keempat dan paling mudah kedua. Selain itu, MSIB di Bangkit Academy memberikan sertifikasi dalam Machine Learning, memberikan nilai tambah bagi mahasiswa dalam mencari pekerjaan. Oleh karena itu, partisipasi dalam program Studi Independen bersama Bangkit Academy dianggap penting, karena tidak hanya memberikan pengetahuan yang berguna tetapi juga memberikan sertifikasi sebagai nilai tambah dalam dunia kerja.

Untuk memperluas pengetahuan dan pengalaman dalam bidang pemrograman Machine Learning, penulis memutuskan untuk melakukan studi independen dengan mengikuti kelas machine learning dalam program Bangkit Academy dan menyelesaikan proyek dengan judul "MOVIEMU". Di era digital

saat ini, jumlah konten film yang tersedia untuk dinikmati konsumen meningkat secara eksponensial. Hal ini menciptakan tantangan bagi pengguna untuk menemukan film yang sesuai dengan preferensi mereka. Salah satu solusi yang dapat digunakan untuk mengatasi tantangan tersebut adalah dengan mengembangkan sistem rekomendasi film yang dapat membantu pengguna menemukan film yang sesuai dengan selera mereka. “MOVIEMU” hadir sebagai solusi permasalahan tersebut dengan mengimplementasikan algoritma cosine kemiripan untuk memberikan rekomendasi film yang akurat dan relevan.

Melalui Proyek “MOVIEMU” bertujuan untuk mengembangkan sistem rekomendasi film yang mampu memberikan rekomendasi yang akurat dan relevan kepada pengguna berdasarkan preferensi mereka. Secara spesifik, proyek ini bertujuan untuk membangun model rekomendasi menggunakan algoritma cosine kemiripan dan mengintegrasikannya ke dalam platform “MOVIEMU”, memastikan sistem berjalan dengan baik dan memberikan rekomendasi yang memuaskan pengguna. Selain itu, proyek ini juga bertujuan untuk menguji sistem rekomendasi, memastikan keakuratan dan relevansi rekomendasi yang diberikan, serta meningkatkan pengalaman pengguna dengan memberikan rekomendasi film yang sesuai dengan preferensi mereka.

Dalam pengerjaan proyek ini penulis sebagai machine learning engineer bertugas untuk membuat model rekomendasi tempat umkm terdekat berdasarkan lokasi dari user. Untuk pembuatan model dalam proyek ini penulis menggunakan metode Feedforward karena memungkinkan model untuk memproses informasi dengan cepat dan efisien.

Algoritma *cosine similarity* adalah salah satu metode populer yang digunakan dalam sistem rekomendasi untuk mengukur kesamaan antara dua item. Algoritma ini bekerja dengan menghitung kosinus sudut antara dua vektor dalam ruang fitur. Dalam konteks rekomendasi film, vektor-vektor ini dapat mewakili fitur-fitur sebuah film, seperti genre, aktor, sutradara, dll. *Cosine similarity* memiliki beberapa keunggulan, antara lain kesederhanaan dan efisiensi, sehingga mudah diterapkan dan tidak memerlukan banyak hal. sumber daya komputasi. Selain itu, algoritma ini mengabaikan besaran vektor dan berfokus pada orientasi vektor sehingga cocok untuk data yang bersifat biner atau kategorikal. Terakhir, *cosine sameness* memiliki skalabilitas yang baik, sehingga dapat diterapkan pada dataset berukuran besar dengan performa yang optimal.

Dalam laporan ini, akan dijabarkan hasil dari studi independen yang telah penulis lakukan selama mengikuti kelas di Bangkit Academy, termasuk dari studi independen yang telah penulis laksanakan selama mengikuti kelas di Bangkit Academy, termasuk pengetahuan yang telah penulis dapatkan, proyek akhir yang telah penulis kerjakan, serta kendala yang penulis hadapi selama proses pembelajaran. Penulis berharap dengan dengan menyelesaikan studi independen pada Bangkit Academy ini dapat mengasah dan memperkuat hard skills maupun soft skills yang terbentuk, sehingga mampu melangkah lebih lanjut dalam dunia industri dan memberikan kontribusi positif pada lingkungan masyarakat.

Metode Pelaksanaan

Proses penyelesaian tugas dalam program pembelajaran ini diselenggarakan secara bervariasi. Setiap akhir sesi ILT (Instructor-Led Training), baik untuk pengembangan soft skills maupun hard skills, mahasiswa diharuskan mengerjakan 10 soal dengan batasan waktu 10 menit. Hasil evaluasi langsung tersedia setelah mahasiswa mengirimkan jawaban mereka. Setiap minggunya, mahasiswa diberikan tugas dengan format yang beragam, termasuk pembuatan esai, video, dan lainnya. Untuk tugas harian dan pemrograman dilakukan secara langsung di platform Coursera, dan nilai untuk setiap tugas juga tersedia secara instan setelah pengumpulan. Project, yang merupakan proyek akhir, dilakukan secara berkelompok dan melibatkan peserta dari berbagai jalur pembelajaran. Proyek ini mengharuskan pembagian tugas yang sesuai dengan jalur pembelajaran masing-masing untuk menghasilkan sebuah aplikasi. Oleh karena itu, mahasiswa membuat timeline proyek agar dapat menyelesaikan tugas dengan tepat waktu. Peyelesaian tugas yang relevan dengan jalur pembelajaran masing-masing menjadi kunci dalam menyelesaikan Project ini dengan sukses. Berikut Timeline penyelesaian proyek ini.

No	Tahapan Pengerjaan Proyek	Periode			
		10 - 12 Mei 2024	12 - 20 Mei 2024	20 Mei - 13 Juni 2024	13 - 22 Juni 2024
1	Business Understanding				
2	Preprocessing Data				
3	Pemodelan				
4	Pengambilan Hasil Akhir				

Hasil dan Pembahasan

Pada Program Studi Independent ini penulis mengembangkan sistem rekomendasi film. Berikut hasil dan pembahasan proyek yang dibagi menjadi 4 proses akan dijelaskan dibawah ini.

1. *Business Understanding*

Algoritma *Cosine Similarity* adalah teknik matematika yang digunakan untuk mengukur kemiripan antara dua objek (biasanya berupa vektor) dalam ruang multidimensi. Dalam konteks sistem rekomendasi film, ini digunakan untuk membandingkan film dan menentukan seberapa

"mirip" mereka satu sama lain.

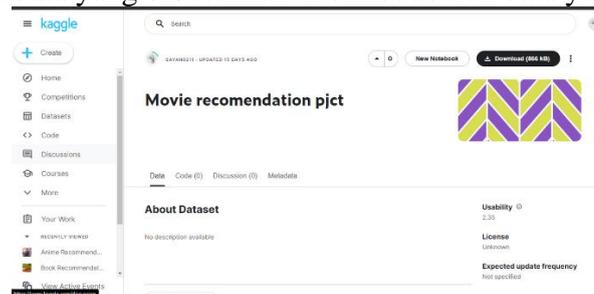
Setiap film dalam sistem rekomendasi diubah menjadi vektor numerik. Setiap elemen dalam vektor mewakili fitur tertentu dari film, seperti kata-kata dalam deskripsi, genre, sutradara, aktor, atau rating pengguna rata-rata. Berdasarkan nilai cosine similarity yang dihitung, sistem merekomendasikan film yang memiliki nilai tertinggi kepada pengguna. Film dengan nilai cosine similarity tertinggi dianggap sebagai film yang paling mirip dengan film yang disukai pengguna sebelumnya.

Dalam Project “**MOVIEMU**”, Algoritma *cosine similarity* adalah metode pembuatan model yang digunakan untuk mengkategorikan film berdasarkan genre. Algoritma ini mampu membedakan genre seperti drama, komedi, aksi, dan sci-fi dengan memahami ciri khas dari masing-masing genre. Hal ini memungkinkan “**MOVIEMU**” untuk merekomendasikan film dari genre yang disukai pengguna di masa lalu, meningkatkan kemungkinan mereka untuk menikmati film yang direkomendasikan.

Algoritma Cosine Similarity, dengan segala kelebihan dan kekurangannya, menjadi fondasi penting dalam sistem rekomendasi film “**MOVIEMU**”. Kemampuannya untuk mengukur kemiripan film berdasarkan konten dan mempersonalisasi rekomendasi berdasarkan pola tontonan Anda menjadikannya alat yang berharga untuk menghadirkan pengalaman menonton yang lebih memuaskan.

2. Preprocessing Data

Data yang digunakan dalam proses ini adalah data yang diambil dari *Kaggle*. Dalam proses ini dilakukan pengambilan data menggunakan *Kaggle* sebagai uji coba secara spesifik sehingga didapatkan data yang sesuai dengan data yang dibutuhkan. Berikut adalah data yang digunakan.



Gambar 2. 1 data yang digunakan dalam project

Data diatas berisi beberapa fitur yang digunakan dalam model, sehingga model dapat membaca data dan menghasilkan rekomendasi dari data yang didapatkan. Berikut penjelasan setiap fitur dari data yang digunakan :

- a. movieId merupakan parameter bernilai unique. Parameter ini digunakan utk mengidentifikasi daftar tiap-tiap film.
- b. userId merupakan parameter bernilai unique. Parameter ini digunakan utk mengidentifikasi daftar tiap-tiap pengguna.
- c. rating merupakan parameter berisi nilai rating film yg diberikan pengguna.

- d. genre merupakan parameter yg menyimpan tiap-tiap kategori film.
- e. title merupakan parameter yg menyimpan judul masing-masing film.

Berikut adalah langkah-langkah yang dilakukan untuk mendapatkan data dan juga membuat model machine learning untuk memberikan rekomendasi tempat. Berikut penjelasan tiap langkahnya.

1. Meload Dataset ke dalam sebuah Dataframe menggunakan pandas
2. df.info() digunakan untuk mengecek tipe kolom pada dataset
3. df.isna().sum() digunakan untuk mengecek apakah ada kolom yg kosong, pada dataset ini nilai kosong tidak ditemukan
4. df.describe() digunakan utk mendapatkan info mengenai dataset terhadap nilai rata-rata, median, banyaknya data, nilai Q1 hingga Q3 dan lain-lain.
5. len(nama_variable.unique())menghitung panjang data unique dari variable tertentu
6. Mengurutkan dataset dan menghapus data movieId yg sama

```

1 # Memuat dataset ke dalam variable baru
2 movie = movies_meta_data.movieId.unique()
3
4 # Mengurutkan data dan menghapus data yang sama
5 movie = np.sort(np.unique(movie))
6
7 print('Jumlah seluruh data movie berdasarkan movieId: ', len(movie))

```

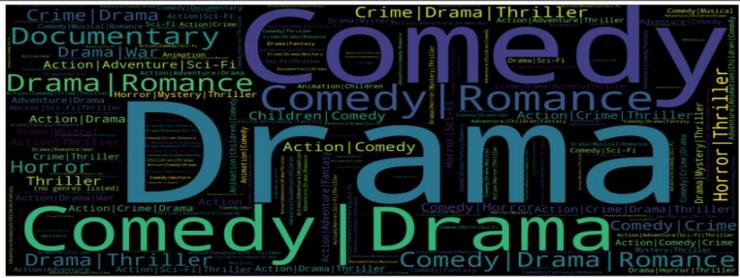
Jumlah seluruh data movie berdasarkan movieId: 9742

7. Menampilkan rata-rata genre yg paling banyak muncul pada dataset

```

1 word_could_dict = Counter(movies_meta_data['genres'].tolist())
2 wordcloud = WordCloud(width = 2000, height = 1000).generate_from_frequencies(word_could_dict)
3
4 plt.figure(figsize=(15,8))
5 plt.imshow(wordcloud)
6 plt.axis("off")
7 plt.show()

```



8. Mereset ulang penomoran indeks data (*data transformation*)

```

1 data.reset_index()
2 data

```

	judul	genre
0	Toy Story (1995)	Adventure Animation Children Comedy Fantasy
1	Jumanji (1995)	Adventure Children Fantasy
2	Grumpier Old Men (1995)	Comedy Romance
3	Waiting to Exhale (1995)	Comedy Drama Romance
4	Father of the Bride Part II (1995)	Comedy
...
9737	Black Butler: Book of the Atlantic (2017)	Action Animation Comedy Fantasy
9738	No Game No Life: Zero (2017)	Animation Comedy Fantasy
9739	Flint (2017)	Drama
9740	Bungo Stray Dogs: Dead Apple (2018)	Action Animation
9741	Andrew Dice Clay: Dice Rules (1991)	Comedy

9737 rows x 2 columns

Teknik yang digunakan pada tahapan Proses Data adalah vektorisasi fungsi CountVectorizer dari library scikit-learn. CountVectorizer digunakan untuk mengubah teks yang diberikan menjadi vektor berdasarkan frekuensi (jumlah) setiap kata yang muncul di seluruh teks. CountVectorizer membuat matriks di mana setiap kata unik diwakili oleh kolom matriks, dan setiap sampel teks dari dokumen adalah baris dalam matriks. Nilai setiap sel tidak lain adalah jumlah kata dalam sampel teks tertentu. Pada proses vektorisasi ini, digunakan metode sebagai berikut.

- *Fit* untuk melakukan perhitungan idf pada data

```

1 # Inisialisasi CountVectorizer
2 tf = CountVectorizer()
3
4 # Melakukan perhitungan idf pada data genre
5 tf.fit(genre)
6
7 # Mapping array dari fitur index integer ke fitur nama
8 tf.get_feature_names_out()

```

```

array(['action', 'adventure', 'animation', 'children', 'comedy', 'crime',
       'documentary', 'drama', 'fantasy', 'fi', 'film', 'genres',
       'horror', 'imax', 'listed', 'musical', 'mystery', 'no', 'noir',
       'romance', 'sci', 'thriller', 'war', 'western'], dtype=object)

```

- *get_feature_names_out* untuk melakukan mapping array dari fitur integer ke fitur nama
- *fit_transform()* untuk mempelajari kosakata dan IDF dengan memberikan nilai return berupa *document-term matrix*

```

[ ] 1 # Melakukan fit lalu ditransformasikan ke bentuk matriks
     2 tfidf_matrix = tf.fit_transform(genre)
     3
     4 # Melihat ukuran matriks tfidf
     5 tfidf_matrix.shape

```

```

(9737, 24)

```

- *Todense()* untuk mengubah vektor TF-IDF dalam bentuk matriks

3. Pemodelan Machine Learning

Setelah dilakukan pra-pemrosesan pada dataset, langkah selanjutnya adalah modeling terhadap data. Pada tahap ini Model machine learning yang digunakan pada sistem rekomendasi ini adalah model content-based filtering dengan simlarty measure yang digunakan adalah Cosine Similarity. Model content-based filtering ini bekerja dengan mempelajari profil minat pengguna baru berdasarkan data dari objek yang telah dinilai pengguna. Metode ini bekerja dengan menyarankan item serupa yang pernah disukai sebelumnya atau sedang dilihat sekarang kepada pengguna berdasarkan kategori tertentu dari item yang dinilai oleh pengguna dengan menggunakan similarity tertentu.

Sedangkan cosine similarity adalah salah satu teknik mengukur kesamaan yang bekerja dengan mengukur kesamaan antara dua vektor dan menentukan apakah kedua vektor tersebut menunjuk ke arah yang sama dengan menghitung sudut cosinus antara dua vektor. Semakin kecil sudut cosinus, semakin besar nilai cosine similarity. cara kerja dari fungsi cosine similiraty yaitu dengan melakukan perhitungan yang sering digunakan untuk menghitung kemiripan diantara item-item. Secara umum, fungsi similarity adalah fungsi yang menerima dua buah obyek berupa bilangan riil (0 dan 1) dan mengembalikan nilai kemiripan (similarity) antara kedua obyek tersebut berupa bilangan riil. Cosine similarity merupakan salah satu metode pengukuran kemiripan yang populer. Metode ini digunakan untuk menghitung nilai kosinus sudut antara dua vektor dan biasanya digunakan untuk mengukur kemiripan antara dua teks/dokumen. Fungsi cosine similarity antara item A dan item B ditunjukkan sebagai berikut.

$$sim(A, B) = \frac{n(A \cap B)}{\sqrt{n(A)n(B)}}$$

Keterangan :

- $Sim(A, B)$ = nilai similaritas dari item A dan item B
- $n(A)$ = banyak fitur konten item A
- $n(B)$ = banyak fitur konten item B
- $n(A \cap B)$ = banyaknya fitur konten yang terdapat pada item A dan B

Berikut cara melatih model dengan menggunakan *cosine similarity* :

```

1 # Menghitung cosine similarity pada matrix tf-idf
2 cosine_sim = cosine_similarity(tfidf_matrix)
3 cosine_sim

array([[1.         , 0.77459667, 0.31622777, ..., 0.         , 0.31622777,
        0.4472136  ],
       [0.77459667, 1.         , 0.         , ..., 0.         , 0.         ,
        0.         ],
       [0.31622777, 0.         , 1.         , ..., 0.         , 0.         ,
        0.70710678],
       ...,
       [0.         , 0.         , 0.         , ..., 1.         , 0.         ,
        0.         ],
       [0.31622777, 0.         , 0.         , ..., 0.         , 1.         ,
        0.         ],
       [0.4472136  , 0.         , 0.70710678, ..., 0.         , 0.         ,
        1.         ]])

```

Pada tahapan ini menampilkan matriks kesamaan setiap judul dengan menampilkan judul film dalam 10 sampel kolom (axis = 1) dan 10 sampel baris (axis=0).

```

1 cosine_sim_df = pd.DataFrame(cosine_sim, index=data['judul'], columns=genre)
2 print("Shape:", cosine_sim_df.shape)
3
4
5 cosine_sim_df.sample(10, axis=1).sample(10, axis=0)

```

Genre	Comedy	Romance	Drama	Musical	Crime	File-Noir	Comedy	Crime	Musical	Crime	Drama	File-Noir	Thriller	Action	Animation	Comedy	Horror	Fantasy	Horror	Documentary	Action	Adventure	Comedy	
judul																								
Dogtown and Z-Boyz (2011)	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000						0.000000				0.0	1.0						
Don't Look Now: We're Being Shot At! La grande vadrouille (1969)	0.500000	0.000000	0.000000	0.500000	0.0	0.000000							0.353553				0.0	0.0						
Natural Born Killers (1994)	0.000000	0.000000	0.333333	0.408248	0.0	0.516398							0.288675				0.0	0.0						
How My Son (1991)	0.707107	0.000000	0.000000	0.707107	0.0	0.000000							0.500000				0.0	0.0						

4. Hasil akhir

Dalam pemanggilan rekomendasi judul film menggunakan function yang dibuat dengan code seperti yg terlihat pada dibawah berikut ini:

```

6.2 Membuat fungsi untuk memanggil 20 rekomendasi film berdasarkan judul yang di input

[35] 1 def movie_recommendations(judul, cosine_sim = cosine_sim, items_data[["judul", "genre"]]):
2     # Mengambil indeks dari judul film yang telah didefinisikan sebelumnya
3     idx = indices[judul]
4
5     # Mengambil skor kemiripan dengan semua judul film
6     sim_scores = list(enumerate(cosine_sim[idx]))
7
8     # Mengurutkan film berdasarkan skor kemiripan
9     sim_scores = sorted(sim_scores, key = lambda x : x[1], reverse = True)
10
11    # Mengambil 20 skor kemiripan dari 1-20 karena urutan 0 memberikan indeks yang sama dengan judul film yang diinput
12    sim_scores = sim_scores[1:20]
13
14    # Mengambil judul film dari skor kemiripan
15    movie_indices = [i[0] for i in sim_scores]
16
17    # Mengembalikan 20 rekomendasi judul film dari kemiripan skor yang telah diurutkan dan menampilkan genre dari 20 rekomendasi film tersebut
18    return pd.DataFrame(data["judul"][movie_indices]).merge(items)

1 # mengecek judul film di dalam data
2 data[data.judul.eq("Johnny English Reborn (2011)")]

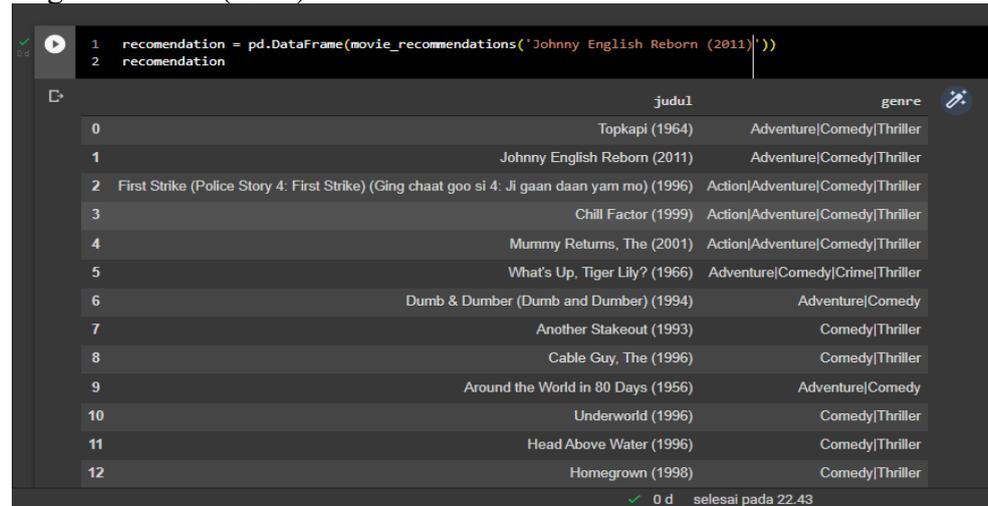
```

Tahapan yang dilakukan pada fungsi tersebut ialah sebagai berikut :

1. Mengambil indeks dari judul film yang telah didefinisikan sebelumnya
2. Mengambil skor kemiripan dengan semua film
3. Mengurutkan film berdasarkan skor kemiripan

4. Mengambil 19 judul berdasarkan kemiripan dari 1-20 karena urutan 0 memberikan indeks yang sama dengan judul film yang diinput
5. Mengambil judul film dari skor kemiripan
6. Mengembalikan 19 rekomendasi judul film dari kemiripan skor yang telah diurutkan dan menampilkan genre dari 19 rekomendasi film tersebut

Berikut top-20 rekomendasi berdasarkan genre dari judul film “Jhonny English Reborn (2011)”



```

1 recommendation = pd.DataFrame(movie_recommendations('Johnny English Reborn (2011)'))
2 recommendation

```

	judul	genre
0	Topkapi (1964)	Adventure Comedy Thriller
1	Johnny English Reborn (2011)	Adventure Comedy Thriller
2	First Strike (Police Story 4: First Strike) (Ging chaat goo si 4: Ji gaan daan yam mo) (1996)	Action Adventure Comedy Thriller
3	Chill Factor (1999)	Action Adventure Comedy Thriller
4	Mummy Returns, The (2001)	Action Adventure Comedy Thriller
5	What's Up, Tiger Lily? (1966)	Adventure Comedy Crime Thriller
6	Dumb & Dumber (Dumb and Dumber) (1994)	Adventure Comedy
7	Another Stakeout (1993)	Comedy Thriller
8	Cable Guy, The (1996)	Comedy Thriller
9	Around the World in 80 Days (1956)	Adventure Comedy
10	Underworld (1996)	Comedy Thriller
11	Head Above Water (1996)	Comedy Thriller
12	Homegrown (1998)	Comedy Thriller

0 d selesai pada 22.43

Kesimpulan

Proyek MOVIEMU program Bangkit Academy bertujuan untuk menciptakan sistem rekomendasi film yang tepat dengan menggunakan algoritma kesamaan kosinus. Metode ini menerapkan efisiensi dan kesederhanaan algoritme untuk membantu pengguna memilih film yang sesuai dengan preferensi mereka. Untuk meningkatkan keterampilan teknis pembelajaran mesin dan memberikan kontribusi yang lebih besar pada industri teknologi, penulis akan berbagi solusi inovatif untuk peningkatan pengalaman pengguna di masa depan.

Ucapan Terimakasih

Saya mengucapkan terima kasih kepada semua pihak yang telah mendukung dan membantu terselesaikannya proyek ini. Terima kasih Bangkit Academy, Google, GoTo, Traveloka, Kementerian Pendidikan dan Kebudayaan RI atas kesempatan dan bimbingannya yang luar biasa. Pengalaman yang tidak bisa dilupakan ini telah memberikan banyak pelajaran berharga. Serta para pembimbing yang telah bersabar dalam memberikan bimbingan dan ilmu. Saya juga mengucapkan terima kasih kepada teman-teman dan rekan-rekan di Bangkit Academy yang selalu mendukung dan memotivasi saya.

Daftar Pustaka

- Merdeka Belajar : Kampus Merdeka - Direktorat Jenderal Pendidikan Tinggi Kementerian Pendidikan dan Kebudayaan.* (2024). [Kemdikbud.go.id](https://kampusmerdeka.kemdikbud.go.id/web/about/latar-belakang).
<https://kampusmerdeka.kemdikbud.go.id/web/about/latar-belakang>
- Abdishakur. (2019, August 4). *Location-Based Recommendations - Towards Data Science*. Medium; Towards Data Science.
<https://towardsdatascience.com/location-based-recommendations->

[bb70af7b1538](#)

Eksplorasi dan Visualisasi Data – eksplorasidata.mipa.ugm.ac.id. (2022).

Ugm.ac.id. <https://eksplorasidata.mipa.ugm.ac.id/eksplorasi-dan-visualisasi-data/>

Feedforward Neural Networks: A Quick Primer for Deep Learning. (2022). Built

In. <https://builtin.com/data-science/feedforward-neural-network-intro>