

Analisa Sistem Rekomendasi Konten Youtube Berdasarkan Durasi Menonton Menggunakan Content-Based Filtering

I Gede Ngurah Wahyu Ananta^{a1}, I Wayan Supriana^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
wahyuananta5@gmail.com
wayan.supriana@unud.ac.id

Abstract

In today's era, the internet is a facility in social life that causes phobias, complete and necessary information is difficult to obtain again. However, how YouTube provides consistently using an algorithm designed for YouTube content recommendations which is an online video media can witness important moments instantly to individuals who are not on television media so that all users can get useful information and entertainment from website media. For some reason Youtube is used as social media with the highest user level from Instagram. Therefore, we make an experiment to categorize the right content to be a crucial factor in producing accurate and meaningful recommendations. In a system analysis, it recommends content on Youtube based on individual categories using the basic concept of the content based filtering algorithm and how it is implemented in the context of YouTube. The model training is carried out using the cosine similarity method which aims to compare the similarity between the contents of these representations. Evaluation of the model can provide insight into how effective the algorithm is in producing relevant recommendations. The steps in the recommendation system analysis are literature study, data collection, model training, and model evaluation by increasing understanding of content-based filtering algorithms.

Keywords: YouTube, content recommendations, content-based filtering, cosine similarity

1. Pendahuluan

Studi pendahuluan dapat memfokus tentang bagaimana perilaku menonton konten di YouTube secara konsisten yang mempengaruhi algoritma untuk rekomendasi konten YouTube yang merupakan media video online dapat menyaksikan momen untuk mencari acara kesukaan secara instan kepada individu yang tidak ada di media televisi yang menjadi semua pengguna bisa didapatkan suatu informasi yang bermanfaat serta hiburan dari media situs web yang berupa video secara bebas tanpa batasan secara langsung. Menurutny dari survei pengguna internet ini bahwa aplikasi yang paling aktif untuk aktivitas berbagai mendominasi adalah Youtube diperoleh sebesar 88% dari jumlah penduduk Indonesia yang aktif menggunakan media sosial. Namun dengan alasan mengapa Youtube yang dijadikan sebagai aplikasi media social yang paling banyak dipakai adalah karena Youtube yang bisa memberikan informasi dan hiburan yang kepada setiap individu dalam bentuk video sehingga Youtube yang menjadi kategori TV dan video kedua yang paling banyak ditelusuri di situs web pada tahun 2020, dengan sebanyak 651 juta pencarian per bulan. Ini mengalahkan Instagram dan kawan-kawan. Dengan demikian, data yang menunjukkan bahwa banyaknya suatu situs yang ditelusuri sehingga data tersebut bisa mendukung konteks media social [1]. Namun ada beberapa alasan yang kebanyakan kalangan senang menikmati sarana Youtube adalah mengenai durasi iklan yang singkat dan sedikit dari televisi, bisa menonton secara on the go, banyaknya pilihan menonton, bisa mengetahui trending, bebas menonton tidak perlu akses akun ataupun berlangganan yang tanpa perlu menunggu atau takut ketinggalan acara, dan bisa berkarya dengan segala kondisi ekonomi. Untuk membangun sebuah topik yang diangkat maka adapun perlu dibahas didalam suatu sistem alur yang bekerja

adalah metode Content-based filtering dalam merekomendasikan konten kemiripan dengan apa yang sering diklik maupun disukai oleh pengguna dengan topik yang sama. Kemudian, data riwayat tersebut otomatis merekam aktivitas user di Youtube yang menyebabkan suatu peristiwa disaat dia sedang menonton dan konten apa yang mereka dinikmati salah satunya adalah melalui mesin proses YouTube Recommendation Algorithm yang bertujuan untuk mengukur parameter di dalam waktu menjalankan platform tersebut artinya jika seseorang seberapa sering menyaksikan atau menikmati yang berbagai topik yang sedang tren lalu dicari atau sekedar langsung diputar saja kemudian adapun faktor dapat dipertimbangkan oleh algoritma tersebut adalah pola kebiasaan yang dilakukan oleh pengguna serupa. Tetapi dengan catatan bahwa platform Youtube tidak bisa secara terbuka yang mengungkapkan seluruh detail mengenai pola algoritma mereka, dan algoritma ini terus berkembang seiring waktunya. YouTube juga mengambil langkah-langkah untuk memastikan bahwa algoritma itu mematuhi kebijakan platform, termasuk kebijakan tentang konten yang aman dan relevan. Berikut adalah beberapa rekomendasi konten YouTube yang berasal dari algoritma content-based filtering adalah seseorang yang berpengalaman sebagai contoh jika seseorang suka menonton video tentang teknologi komputer, maka yang dijalankan oleh algoritma content-based filtering yang akan merekomendasikan video tentang teknologi lainnya, seperti video tentang gadget terbaru atau tutorial penggunaan software yang tertentu melalui tata letak pada aplikasi tersebut baik di menu beranda maupun scrolling dibawahnya namun tetapi jika dia sering menonton tentang musik, maka algoritma content-based filtering yang bekerja didalam suatu pemrosesan yang akan merekomendasikan video tentang musik lainnya, seperti video klip musik atau konser musik. Metode content-based filtering yang mempertimbangkan fitur-fitur dari konten yang disukai oleh pengguna, seperti tag, deskripsi, dan judul, namun untuk menentukan seberapa jawaban yang muncul dapat dikatakan konten untuk ditonton berikutan dengan isi yang relevan berdasarkan pola mengklik video yang secara terus menerus. Dengan demikian, metode ini memiliki kelebihan dalam merekomendasikan konten yang spesifik dan personal ada juga memiliki kelemahan dalam merekomendasikan konten yang berbeda dari minat pengguna.

Algoritma content-based filtering pada YouTube menentukan rekomendasi video dengan mempertimbangkan fitur-fitur dari konten yang disukai oleh pengguna, seperti tag, deskripsi, dan judul. Algoritma ini kemudian mencari video lain yang memiliki fitur-fitur serupa dengan konten yang disukai oleh pengguna menggunakan metode cosine similarity. Metode cosine similarity digunakan untuk mengukur seberapa mirip dua vektor dalam ruang fitur yang diterapkan pada aplikasi berbasis mobile dengan menggunakan persentase atau data yang terukur terhadap kemiripan produk dengan hasil kemiripan yang tertinggi [3]. Semakin tinggi nilai cosine similarity, semakin mirip kedua vektor tersebut. Dengan demikian, semakin tinggi nilai cosine, semakin mirip konten kedua video tersebut dan semakin besar kemungkinan video tersebut direkomendasikan kepada pengguna. Algoritma content-based filtering pada YouTube juga mempertimbangkan riwayat tontonan dan perilaku pengguna dalam menonton video sebelumnya, seperti rewind dan rewatch, untuk merekomendasikan video yang sesuai dengan minat dan riwayat pengguna.

2. Metode Penelitian

Pada bagian ini yang dilakukan beberapa penelitian sebelumnya yaitu analisa e-commerce namun dilakukan dengan studi kasus berbeda terkait analisa konten di Youtube adalah sebagai berikut :

2.1 Studi Literatur

Untuk suatu pada algoritma content-based filtering di Youtube bahwa alur pada rekomendasi konten yang melibatkan pengguna suatu kumpulan informasi dalam konten youtube meliputi tanggal, judul, channel, kategori, atau durasi yang bisa membuat rekomendasi terbaik yang personal. Algoritma ini memiliki mengukur kemiripan konten yang bernama cosine similarity [3]. Hasil penelitian nantinya bahwa bisa menunjukkan pada Content-Based Filtering yang efektif dari durasi menonton dalam menghasilkan rekomendasi relevan berdasarkan preferensi pengguna dan karakteristik video.

2.2. Pengumpulan Data

Dalam suatu sistem ini yang dapat mengumpulkan topik video yang ada di Youtube tentu bisa digunakan untuk mengetahui informasi tentang merekomendasi konten. Kemudian data ini harus mencakup pada informasi seperti judul video, nama channel, tag, kategori, dan informasi usia terkait video tersebut. mengidentifikasi saat simulasi aplikasi Youtube berjalan dengan pencarian atau mengklik langsung dari tampilan beranda dan menyimpan data history tersebut secara sembunyi khusus untuk mode samaran dalam percobaan langsung.

2.3. Pelatihan Model

Setelah melakukan pemodelan, maka data dapat dilatih oleh model machine learning video berdasarkan pencarian judul video tersebut yang diimbangi dengan kategori ke suatu algoritma yang dapat bekerja dalam pembelajaran mesin seperti collaborative filtering, content-based filtering, atau hybrid filtering yang bertujuan untuk menghasilkan suatu rekomendasi. Dalam suatu implementasinya maka proses ini akan membandingkan minat dengan pengguna produk termasuk platform digital lalu untuk mengoptimalkan parameter model dalam proses mencapai kinerja terbaik yang menggunakan tools cosine similarity.

2.4. Pengujian Model

Setelah kita pelatihan berdasarkan judul video maka akan diuji dari beberapa kumpulan konten tersebut yang dipakai dalam mengolah suatu rekomendasi konten. Dari suatu pengujian sistem yang akan membantu dalam memvalidasi yang terkait dengan keefektifan algoritma, penyesuaian yang dilakukan, serta meningkatkan pengalaman pengguna didalam menemukan suatu konten yang relevan dari Youtube.

2.5 Evaluasi Model

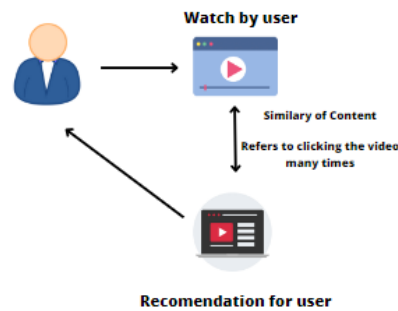
Dengan adanya evaluasi model ini maka diperlukan menguji dalam suatu kinerja pemodelan. Kemudian, ini dapat membagikan data konten Youtube yang menjadi antara set pelatit dan pengujian. Adapun yang dalam evaluasi model ini berdasarkan waktu menonton yang dapat diberikan sinyalir kepada suatu sistem rekomendasi mengacu ke pengguna. Untuk memastikan bahwa penonton dapat dikatakan puas dengan konten yang akan ditonton, sistem tersebut dapat mengukur tiap konten video tersebut [4]. untuk mengevaluasi model ini, saya menggunakan metrik evaluasi precision dengan rumus dasarnya adalah:

$$\text{recommender system precision: } P = \frac{\# \text{ of our recommendations that are relevant}}{\# \text{ of items we recommended}} \quad (1)$$

2.6 Perancangan Sistem Rekomendasi

Penelitian ini telah melakukan tahapan-tahapan dalam pengolahan data dimana sebuah sistem rekomendasi ini yang menggunakan algoritma Content-Based Filtering berdasarkan kata kunci di judul video tersebut serta durasi menontonnya. Maksud dari sistem rekomendasi ini untuk memberikan saran dalam proses pengambilan keputusan tiap item. Berikut gambar secara umum dari algoritma Content-Based Filtering

CONTENT-BASED FILTERING

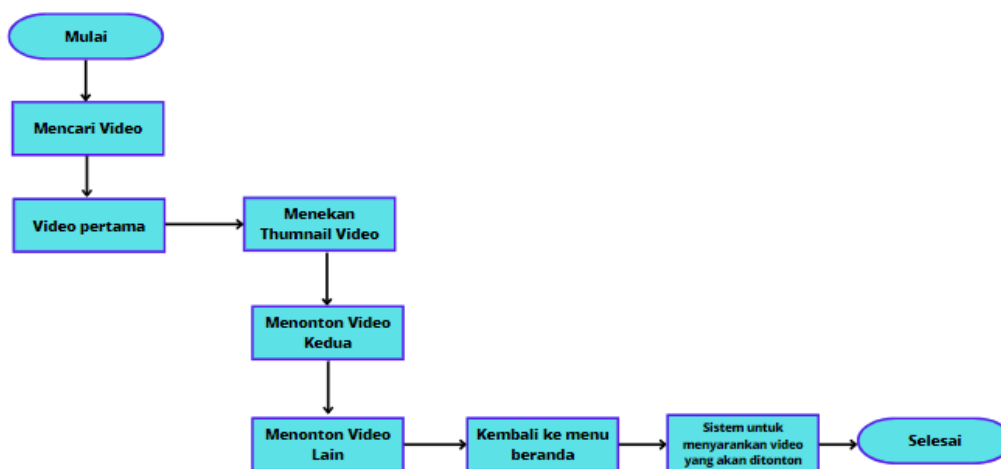


Gambar 1. Alur Content Based Filtering

Lalu, disini pada gambar diatas mengenai content based filtering ini yang membahas bagaimana suatu alur secara umum pada gambar diatas ini yang dirancang dalam suatu rekomendasi konten berupa media video online di Youtube. Implementasi algoritma ini melibatkan perbandingan minat pengguna sesuai data riwayat tontonan yang melalui faktor-faktor diperhitungkan yang berasal dari user sendiri yaitu seberapa sering mengklik suatu video dan menambahkan ke daftar favorite apabila akun tersebut telah dilogin. Namun, sebelum membangun sistem ini dapat diutamakan adalah mengumpulkan perfernsi suatu pengguna terhadap konten kesenangan mereka. Profil pengguna yang dibangun berdasarkan pilihan konten, preferensi, dan selera mereka, yang kemudian ini digunakan untuk lakukan peringkat pengguna. Jadi, algoritma tersebut pada dasarnya mencari kesamaan antara konten-konten yang ada. Kemiripan dan kedekatan ini antara konten-konten diukur berdasarkan kesamaan isi, seperti genre, kategori, dan aspek-aspek lainnya yang dalam konten tersebut yang diharapkan bisa memberikan saran secara singkat sesuai dengan minat dan preferensi para pengguna. Dalam konteks tersebut ini bisa membantu pengguna menemukan konten yang relevan serta menarik berdasarkan suatu kesamaan isi dari konten yang telah mereka dinikmati.

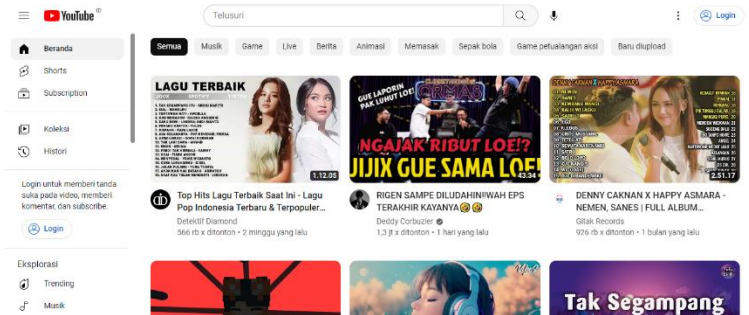
3. Hasil dan Pembahasan

Berikut adalah sebuah gambaran flowchart yang digunakan dalam menganalisa perancangan sistem untuk user melihat beberapa video yang ingin ditonton, saat ditonton, dan seterusnya maka yang akan keluar adalah iklan atau rekomendasi video berikutnya akan ditonton disaat nanti pada menu beranda :



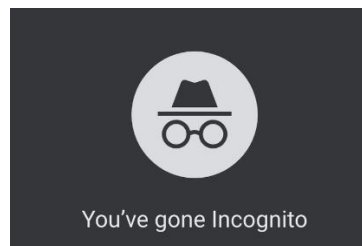
Gambar 2. Flowchart

Kemudian, untuk membuktikan suatu algoritma yang bekerja di platform digital Youtube maka yang kami akan lakukan dengan percobaan dalam mensimulasikan yang menggunakan mode samaran. Ini tentu bertujuan untuk mensetting awal layaknya akun baru pada tampilan beranda yang bersifat sementara. Berikut adalah hasil dokumentasinya :



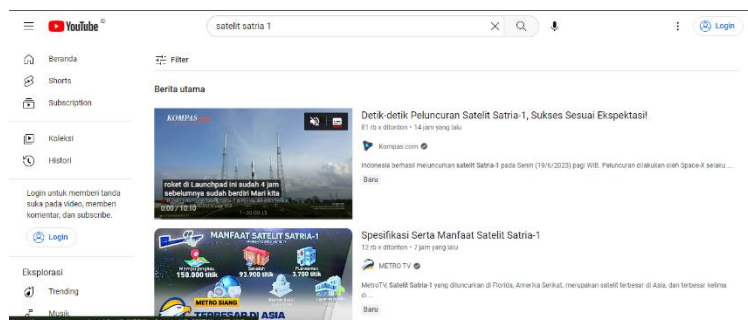
Gambar 2. Flowchart

- a. User dapat membuka aplikasi atau website Youtube yang tersedia. Kemudian, ketika diproses oleh sistem sampai di tampilan beranda dan pastikan dalam keadaan mode samaran. Untuk pada tampilan awal atau beranda anda dapat melihat daftar konten.



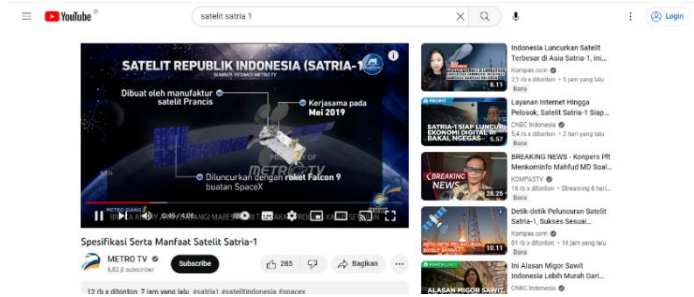
Gambar 2. Flowchart

- b. Mode samaran ini berfungsi untuk menampilkan gambar miniature video atau judul-judul video secara singkat pada halaman beranda. Selain itu, mode ini dapat membantu dalam mengeksplorasi dan menemukan konten video yang menarik tanpa perlu membuka setiap video secara penuh.



Gambar 2. Flowchart

- c. Untuk mencari video yang spesifik, maka user dengan mengklik video simbol pencarian yang terletak di sebelah kanan atas layar. Setelah ini anda cukup mengetikkan topik yang ingin dicari setelah dikonfirmasi maka akan muncul daftar video dengan topik yang ingin menelusuri.



Gambar 2. Flowchart

- d. Setelah menemukan video yang diinginkan dari hasil mesin penelusuran terhadap suatu konten maka user dapat mengklik langsung ke suatu video untuk dapat diputar. Jika setelah itu, anda dapat menonton sebagian selama 3 menit sesuai kebutuhan dari total keseluruhan video yang dimiliki oleh satu judul video. Setelah itu, anda dapat memilih satu video lagi untuk melanjutkan menonton video berikutnya atau tetap pada video yang sama.

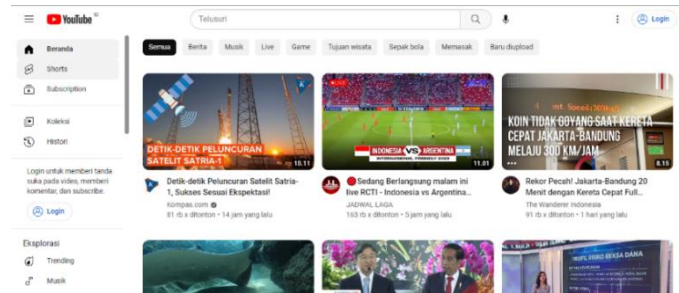


Gambar 2. Flowchart

- e. Lalu, jika user dapat menggantikan video berikutnya dengan topik yang sama maupun lainnya. Anda dapat mencari video baru atau menggunakan fitur rekomendasi secara otomatis pada saat menikmati konten yang diinginkan di samping yang sedang diputar.
- f. Setelah menonton 5 video dengan durasi bervariasi dalam berselang waktu sekitar 10 menit, maka pengguna bisa kembali untuk beralih ke tampilan beranda. Hal ini dapat memberikan pengalaman total waktu menonton yang dapat mempengaruhi rekomendasi konten video dari sistem tersebut.



Gambar 2. Flowchart



Gambar 2. Flowchart

- g. Setelah 10 menit dijeda, maka sistem bisa menunjukan rekomendasi konten yang serupa dan relevan dengan topik yang telah anda tonton sebelumnya di dalam preview tersebut. Semakin lama menonton serta berinteraksi dengan konten di Youtube maka yang muncul terjadi adalah banyak rekomendasi mengenai konten berhubungan dengan topik-topik tertentu di halaman beranda. Dengan mengikuti langkah ini dimana web/aplikasi Youtube ini melakukan simulasi pengalaman dengan mencari, menonton, dan menjelajahi berbagai video yang sesuai dengan minat dan kebutuhan anda.

Adapun sistem rekomendasi dari konten Youtube menggunakan parameter dan algoritma untuk bekerja secara otomatis. Lalu beberapa parameter yang digunakan dalam sistem rekomendasi bekerja konten di Youtube dengan mempertimbangkan beberapa faktor yaitu Riwayat menonton, preferensi pengguna, interaksi pengguna, konteks video, kesesuaian demografis, popularitas trending, pengelompokan konten, durasi tontonan, dan fitur rekomendasi tambahan. Namun, perlu untuk diperhatikan bahwa Youtube terus mengembangkan dan memperbaiki algoritma rekomendasi mereka serta parameter yang dijadikan faktor digunakan dalam sistem rekomendasi dapat berubah seiring dengan waktu.

Kemudian, dapat dipastikan dengan percobaan cosine simily yang dapat dikatakan untuk mengukur kemiripan. Pada sebelumnya pengukuran ini dengan biasa digunakan dalam penambangan teks, pencarian informasi, dan aplikasi pembelajaran mesin seperti klasifikasi dokumen, sistem rekomendasi, dan pengelompokan lalu berikut pada tabelnya:

Title	Duration (Minutes)	Total Similarity
Spesifikasi Serta Manfaat Satelit Satria-1	4	1. 0.40824829 0.32732684 0.36927447 0.35355339
Layanan Internet Hingga Pelosok, Satelit Satria-1 Siap Mengorbit	6	0.40824829 1. 0.4454354 0.40201513 0.38490018
Satelit Satria-1 Diluncurkan Ke Orbit, Dapat Percepat Akses Layanan Internet Merata Se-Indonesia	1	0.32732684 0.4454354 1. 0.40291148 0.46291005
Peluncuran Satelit Satria-1 Sukses, Internet Di Indonesia Akan Semakin Cepat	3	0.36927447 0.40201513 0.40291148 1. 0.34815531
Satelit Satria-1 Diluncurkan, Kini Desa-Desa Bisa Dapat Internet	3	0.35355339 0.38490018 0.46291005 0.34815531 1.

Dari data tersebut melalui tahap pelatihan permodelan yang menyatakan bahwa terdapat penyesuaian yang dilakukan untuk pengukuran dengan menggunakan metode cosine similarity yang memberikan hasil dari pengklasifikasi konten dari masing-masing judul video yang berbeda maupun kata-kata yang sama.

4. Kesimpulan

Pada simulasi yang menjalankan aplikasi Youtube secara langsung dalam analisa sebuah sistem yang berjalan yang mengetahui umpan balik pada pilihan video tersebut di menu beranda telah diterapkan untuk memastikan privasi pengguna dan memberi pengalaman disaat menonton. Dengan menganalisis durasi dan judul video yang ditonton oleh pengguna, arsitektur program ini menggunakan fungsi tools/library seperti CountVectorizer dan cosine_similarity dari perpustakaan sklearn untuk menghitung skor kemiripan yang mengetahui bagaimana cara kerja merekomendasikan konten yang relevan berdasarkan preferensi pengguna. Sistem yang diterapkan menggunakan content-based filtering untuk merekomendasikan video YouTube. Perpaduan antara judul dan lama menontonnya maka yang disebut atribut dalam analisis pengklasifikasian konten disaat berjalan. Dengan memasukkan umpan balik pengguna, pembuatan profil pengguna, pemfilteran kolaboratif, dan pembaruan waktu nyata, sistem dapat lebih ditingkatkan dalam jangka waktu yang belum pasti untuk memberikan rekomendasi yang lebih akurat dan dipersonalisasi dilakukan pengukuran antara judul video dengan durasi menonton untuk meningkatkan pengalaman pengguna secara keseluruhan di platform YouTube.

Daftar Pustaka

- [1] Jati, Nala Prasasta; ABIDIN, Zainal; Oxygentri, Mickey. Pengaruh Sosial Media Youtube Terhadap Perilaku Literasi Siswa SMA Utama. *Linimasa: Jurnal Ilmu Komunikasi*, 2022, 5.1: 67-75.
- [2] Rahmadi, Ridho, et al. Analisis Sentimen Politik Berdasarkan Big Data dari Media Sosial Youtube: Sebuah Tinjauan Literatur. *Automata*, 2021, 2.1.
- [3] Sujasman, M. Bagus; Diana, Diana; SYAZILI, Ahmad. Implementasi Metode Cosine Similarity Untuk Rekomendasi Produk Pada Aplikasi Penjualan Berbasis Mobile. In: *Bina Darma Conference on Computer Science (BDCCS)*. 2020. p. 162-171.
- [4] Cristos Goodrow. "indonesia.googleblog.com" : Memahami cara kerja sistem rekomendasi Youtube. 18 10 2021. [Online]. Available: <https://indonesia.googleblog.com/2021/10/memahami-cara-kerja-sistem-rekomendasi.html?m=1>. [Accessed 9 06 2023].
- [5] Fitrianti, Betharia Sri; Fachurrozi, Muhammad; Yusliani, Novi. Sistem Rekomendasi Artikel Ilmiah Berbasis Web Menggunakan Content-based Learning dan Collaborative Filtering. *Generic*, 2018, 10.1: 17-23.