

**SISTEM REKOMENDASI LAGU TERATAS – GLOBAL
BERDASARKAN PROPERTI TREK MENGGUNAKAN METODE
CONTENT-BASED FILTERING**

LAPORAN TUGAS AKHIR

Laporan ini dibuat untuk memenuhi kelulusan Matakuliah Program Tugas Akhir



Dibuat Oleh:

1.18.4.018

Naomi C.H Tampubolon

**PROGRAM DIPLOMA IV TEKNIK INFORMATIKA
POLITEKNIK POS INDONESIA
BANDUNG
2022**

***TOP SONG RECOMMENDATION SYSTEM – GLOBAL BASED ON
TRACK PROPERTY USING METHOD CONTENT-BASED FILTERING***

FINAL REPORT

This report was created as the requirement for passing the final project course.



Created By:

1.18.4.018

Naomi C.H Tampubolon

DIPLOMA IV PROGRAM IN INFORMATION ENGINEERING

POLITEKNIK POS INDONESIA

BANDUNG

2022

LEMBAR PENGESAHAN

SISTEM REKOMENDASI LAGU TERATAS GLOBAL BERDASARKAN PROPERTI TREK PENGELOMPOKAN MENGGUNAKAN METODE COLLABORATIVE FILTERING

Naomi C.H Tampubolon

1.18.4.018

Laporan Program Tugas Akhir ini telah diperiksa, disetujui dan disidangkan
Di Bandung, 30 Agustus 2022
Oleh :

Ketua Pengaji



Anggota Pengaji



Rolly Maulana Awangga, S.T.,

M.T.,CAIP.,SFPC

NIK. 117.86.219

Nisa Hanum Harani, S.Kom.,

M.T.,CDSP.,SFPC

NIK. 117.89.223

Menyetujui,
Koordinator Tugas Akhir



Roni Andarsyah, ST., M.KOM.,SFPC

NIK. 115.88.193

6058d6da-5469-4fca-b998-3cf171c7654f

LEMBAR PENGESAHAN

SISTEM REKOMENDASI LAGU TERATAS GLOBAL BERDASARKAN PROPERTI TREK PENGELOMPOKAN MENGGUNAKAN METODE COLLABORATIVE FILTERING

Naomi C.H Tampubolon

1.18.4.018

Laporan Program Tugas Akhir ini telah diperiksa, disetujui dan disidangkan
Di Bandung, 30 Agustus 2022

Oleh :

Pembimbing Utama,



Pembimbing Pendamping



Rolly Maulana Awangga, S.T.,

Noviana Riza, S.Si., M.T.,SFPC

M.T.,CAIP.,SFPC

NIK. 103.78.065

NIK. 117.86.219

Menyetujui,

Ketua Program Studi D4 Teknik Informatika,



M. Yusril Helmi S, S.Kom., M.Kom.

NIK. 113.74.163

ABSTRAK

Spotify saat ini merupakan platform musik, podcast dan video yang banyak digemari banyak penggunanya. Spotify juga memberikan kemudahan akses musik. Platform musik spotify sangat banyak digemari penggunanya, terutama kalangan young adults yaitu berusia 18 hingga 25 tahun.(Wahyu Dewatara & Monik Agustin, 2019) Spotify menyediakan banyak fitur seperti yaitu sekarang memutar, cari, koleksi, dengarkan offline, podcast dan acara serta banyak lagi. Properti trek di Spotify termasuk genre, nama artis, dan banyak lainnya. (Fitriani, n.d.) Berdasarkan properti trek, bagaimana cara merekomendasikan lagu berdasarkan properti trek, dan bagaimana cara memastikan bahwa lagu berdasarkan properti trek disukai. Masalah ini akan menjadi sumber dalam penelitian. *Dataset* yang digunakan dalam penelitian ini adalah 50 Lagu Teratas – Global (*50 Top Songs - Global*) yang diambil dari Spotify playlist chart yang kemudian diconvert ke dalam dalam *file* dengan *format cvs* dan kemudian nantinya akan dikelola atau diproses. Penelitian ini bertujuan untuk rekomendasikan lagu berdasarkan trek properti yang ada dan rekomendasikan berdasarkan trek properti disukai. Metode Content-Based Filtering digunakan dalam penelitian ini. *Content-Based Filtering* berfungsi merekomendasikan berdasarkan kemiripan konten seperti *genres* yang merupakan kemiripan jarak *cosinusnya* dan terdapat 5 lagu teratas yaitu *STAYING ALIVE (feat. Drake & Lil Baby)*, *Bam Bam (feat. Ed Sheeran)*, *About Damn Time*, *Bad Decisions (with BTS & Snoop Dogg)*, dan *Ferrari* yang direkomendasikan menggunakan metode *Content-Based Filtering*.

Kata Kunci: Sistem rekomendasi, musik, *spotify*, *content-based filtering*.

ABSTRACT

Spotify is currently a music, podcast and video platform that is loved by many users. Spotify also provides easy access to music. Spotify's music platform is very popular with users, especially young adults aged 18 to 25 years. (Wahyu Dewatara & Monik Agustin, 2019) Spotify provides many features such as now playing, searching, collecting, listening offline, podcasts and shows and many more . Spotify also has track properties such as genre, artist name, and more.(Fitriani, n.d.) Based on the track properties, how to recommend a song based on track properties and how songs recommended based on recommended track properties are liked. These problems will be the material in this research. The dataset used in this study is the Top 50 Songs – Global (50 Top Songs - Global) which are taken from the Spotify playlist chart which is then converted into a file with cvs format and then later it will be managed or processed. This study aims to recommend songs based on existing track properties and recommended based on preferred track properties. In this study using the method of Content Based Filtering. Content Based Filtering serves to recommend based on the similarity of content such as genres which are the similarity of the cosine distance and there are 5 top songs, namely STAYING ALIVE (feat. Drake & Lil Baby), Bam Bam (feat. Ed Sheeran), About Damn Time, Bad Decisions (with BTS) & Snoop Dogg), and a recommended Ferrari using the Content-Based Filtering method.

Keywords: Recommendation system, music, spotify, content-based filtering.

KATA PENGANTAR

Puji dan Syukur kita panjatkan ke hadirat Tuhan Yang Maha Esa, karena telah melimpahkan berkat dan rahmat-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul Sistem Rekomendasi Teratas – Global berdasarkan Properti Trek menggunakan Metode *Content-based filtering* ini dapat diselesaikan dengan baik adanya sebagai syarat kelulusan matakuliah Tugas Akhir.

Namun keberhasilan penulisan laporan ini bukan hanya semata usaha penulis saja, tapi juga banyak bantuan dukungan dari orang-orang sekitar. Dan pada kesempatan ini penulis ingin mengucapkan terimakasih kepada beliau-beliau atau semua pihak yang sudah membantu penyusunan laporan Tugas Akhir ini. Diantaranya:

1. Bapak Rolly Maulana Awangga,S.T.,MT.,CAIP, SFPC. Selaku bapak dosen Pembimbing Utama dan Ibu Noviana Riza,. S.Si., M.T.,SFPC selaku Ibu dosen Pembimbing Pendamping Tugas Akhir.
2. Ibu Nisa Hanum Harani, S.Kom., M.T.,CDSP, SFPC. Selaku ibu dosen Pengaji Tugas Akhir
3. Orangtua kami yang senantiasa mendoakan dan kelancaran pengususunan laporan dan sidang Tugas Akhir.
4. Bapak Roni Andarsyah, S.T., M.Kom., SFPC selaku bapak koordinator Tugas Akhir.
5. Bapak Mohamad Nurkamal Fauzan, S.T., M.T., SFPC selaku Wali Kelas 4B D4 TI.
6. Bapak M. Yusril Helmi Setyawan, S.Kom., M.Kom.,SFPC. selaku Ketua Program Studi DIV Teknik Informatika.
7. Teman seperjuangan YO DREAM dan S U * R A.
8. Dan kepada EXO, NCT, NCT Dream, NCT 127, dan WayV yang telah menemani dan sebagai penyemangat penulis melalui lagu yang mereka sumbangkan dalam pelaksanaan laporan Tugas Akhir ini.

Penulis memohon kepada Pembaca, apabila menemukan kesalahan ataupun kekurangan dalam penulisan Laporan Tugas Akhir ini, dari segi Bahasa maupun Isi, sekiranya pembaca memahami bahwa buku ini masih jauh dari kata sempurna, sehingga penulis sangat mengharapkan kritik serta saran yang bersifat membangun demi terciptanya buku yang akan penulis susun selanjutnya menjadi lebih baik lagi adanya. Mohon maaf jika masih banyak

kekurangan dan semoga laporan ini memberi manfaat untuk setiap pembaca dan juga menambah ilmu bagi penulis. Terimakasih.

Bandung, 30 Agustus 2022

Naomi C.H Tampubolon

1184018

DAFTAR ISI

ABSTRAK	iv
<i>ABSTRACT</i>	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xi
BAB I PENDAHULUAN.....	I-1
1.1 Latar Belakang	I-1
1.2 Identifikasi Masalah	I-2
1.3 Tujuan dan Manfaat	I-2
1.4 Batasan Masalah.....	I-2
1.5 Sistematika Penulisan.....	I-3
BAB II TINJAUAN PUSTAKA	II-5
2.1 State of The Art (SOTA).....	II-5
2.2 Tinjauan Pustaka	II-6
2.3 Landasan Teori	II-12
2.3.1 <i>Spotify</i>	II-12
2.3.2 <i>Recommendation System</i> (Sistem Rekomendasi)	II-15
2.3.3 <i>Content-Based Filtering</i>	II-15
2.4 Kajian Pendekatan yang diusulkan	II-16
2.5 Karakteristik Data yang digunakan	II-16
BAB III METODOLOGI PENELITIAN	III-18
3.1 Diagram Alur Metodologi Penelitian.....	III-18
3.2 Indikator Capaian	III-18
3.3 Tahapan-Tahapan Metode Penelitian.....	III-20

3.3.1	Indentifikasi Masalah.....	III-20
3.3.2	Studi Literatur	III-20
3.3.3	Pengumpulan Data	III-20
3.3.4	Proses Data.....	III-21
3.3.5	Pemodelan dan Pengujian.....	III-21
3.3.6	Hasil Evaluasi	III-22
3.3.7	Diseminasi Hasil	III-22
BAB IV PEMBAHASAN DAN HASIL		IV-23
4.1	Pembahasan	IV-23
4.2	Hasil	IV-38
BAB V PENUTUP		V-39
5.1.	Kesimpulan.....	V-39
5.2.	Saran.....	V-39
DAFTAR PUSTAKA		i
LAMPIRAN-LAMPIRAN		iv

DAFTAR GAMBAR

<i>Gambar II- 1 Alur Content-based filtering.....</i>	<i>II-15</i>
<i>Gambar III- 1 Diagram Alur Metodologi Penelitian</i>	<i>III-18</i>
<i>Gambar III- 2 Flowchart pengumpulan data</i>	<i>III-21</i>
<i>Gambar IV- 1 Playlist spotify lagu teratas -global</i>	<i>IV-24</i>
<i>Gambar IV- 2 Import Library.....</i>	<i>IV-34</i>
<i>Gambar IV- 3 Menampilkan dataset.....</i>	<i>IV-34</i>
<i>Gambar IV- 4 Drop tabel.....</i>	<i>IV-35</i>
<i>Gambar IV- 5 Missing Value.....</i>	<i>IV-35</i>
<i>Gambar IV- 6 Mengisi nilai NA/Nan</i>	<i>IV-35</i>
<i>Gambar IV- 7 Drop entri.....</i>	<i>IV-35</i>
<i>Gambar IV- 8 Encode Data</i>	<i>IV-36</i>
<i>Gambar IV- 9 Rekomendasi lagu.....</i>	<i>IV-37</i>
<i>Gambar IV- 10 Pemodelan dan Pengujian data menggunakan Conten-Based Filtering ..</i>	<i>IV-37</i>

DAFTAR TABEL

<i>Tabel II- 1 Tinjauan Pustaka</i>	<i>II-6</i>
<i>Tabel II- 2 Karakteristik Data</i>	<i>II-16</i>
<i>Tabel III- 1 Indikator Capaian</i>	<i>III-18</i>
<i>Tabel IV- 1 Playlist spotify berbentuk file csv</i>	<i>IV-25</i>
<i>Tabel IV- 2 Hasil.....</i>	<i>IV-38</i>

BAB I

PENDAHULUAN

1.1 Latar Belakang

Musik terdiri dari beberapa unsur, seperti melodi, harmoni, ritme, dan timbre. Musik adalah fenomena *intuitif* yang dapat dikembangkan, ditingkatkan, dan disebarluaskan dengan berbagai cara. Musik adalah fenomena unik yang dihasilkan oleh berbagai alat musik. Musik memang menjadi salah satu hobi yang paling digemari, hal tersebut terbukti dengan manfaat yang bisa didapatkan dengan mendengarkan musik. Musik memiliki kemampuan untuk mengurangi stres. Manfaat mendengarkan musik adalah kita bisa mendengarkannya sambil melakukan hal lain. Musik adalah komponen yang tak terpisahkan dari kehidupan dan teknologi. Perkembangan yang semakin berkembang pesat juga mempengaruhi dunia industri.(Muslimin M, n.d.) Dimana pada jaman era digital saat ini tidak perlu repot dalam mengengarkan musik dimanapun dan kapanpun, dikarenakan kecanggihan teknologi tersebut. *Spotify* adalah layanan *streaming* musik terpopuler di Indonesia, telah mencapai 200 juta pengguna perbulan.(Noviani et al., 2020) Data yang digunakan pada penelitian ini adalah data lagu yang diambil dari *playlist chart spotify 50 Lagu Teratas – Global (50 Top Songs - Global)*. (Fitriani, n.d.) *Spotify* memiliki properti lagu yang berfungsi untuk mengetahui lebih mendalam mengenai lagu atau sebuah musik. Ada beberapa fitur atau properti lagu pada *spotify* yaitu, sebagai berikut: *track name, album name, artist name, duration, popularity, genres, danceability, energi, key, loudness, speechiness, acousticness, instrumentaness, liveness, tempo* dan *valence*. Berdasarkan properti trek tersebut kita dibutuhkan sistem yang dapat merekomendasikan berdasarkan properti trek yang ada. Maka dari itu bagaimana caranya untuk merekomendasikan sebuah lagu berdasarkan properti trek yang ada. Dari permasalahan tersebut dibutuhkan sebuah sistem yang dapat merekomendasikan lagu berdasarkan properti trek yang ada. Properti trek ‘*genres*’ merupakan properti trek yang digunakan dalam rekomendasi sistem.

Dalam penelitian ini sistem rekomendasi akan dianalisa dan dibangun dengan menggunakan metode *content-based filtering* karena dengan algoritma ini berdasarkan *similarity* konten sehingga rekomendasi yang dihasilkan mempunyai kualitas yang baik. Visualisasi hasil

rekomendasi tersebut akan diimplementasikan dengan *framework Jupyter Notebook*. Target luaran yang ingin dicapai pada penelitian ini adalah jurnal nasional terakreditas.

1.2 Identifikasi Masalah

Adapun identifikasi masalah pada penelitian ini sebagai berikut:

- a) Bagaimana menganalisa data 50 lagu teratas-global?
- b) Bagaimana membuat model sistem rekomendasi lagu berdasarkan properti trek ‘genres’?
- c) Bagaimana mengimplementasikan proses data yang sudah diuji dalam sebuah sistem rekomendasi untuk hasil pengujian lagu?

1.3 Tujuan dan Manfaat

Adapun tujuan pada penelitian ini sebagai berikut:

- a) Untuk memberikan rekomendasi lagu berdasarkan properti trek ‘genres’ pada sebuah sistem yang diberikan.
- b) Memberikan pengalaman baru tentang rekomendasi atau *mengeksplor* lagu-lagu yang mungkin belum pernah didengarkan atau disukai.
- c) Menggunakan pendekatan sistem rekomendasi yaitu model *content-based filtering* untuk pemodelan sistem rekomendasi lagu berdasarkan properti trek lagu.
- d) Menerapkan *framework Jupyter Notebook* untuk menampilkan hasil rekomendasi sistem.

Adapun manfaat pada penelitian ini sebagai berikut,

- a) Merekendasikan lagu-lagu berdasarkan perproperti trek.
- b) Membuat tampilan *framework* agar mudah digunakan untuk melakukan proses perekendasian.

1.4 Batasan Masalah

Adapun ruang lingkup pada penelitian ini sebagai berikut,

- a) Dikarenakan kondisi pandemi, data yang digunakan dalam penelitian ini bersumber dari data yang di *convert* dari *playlist chart* pada *Spotify*.

- b) Periode waktu 1 minggu.
- c) Bahasa pemrograman menggunakan *python*.
- d) *Software* yang digunakan *Jupyter* dan *Anaconda*.

1.5 Sistematika Penulisan

Berdasarkan latar belakang dan perumusan masalah diatas, maka penyusunan laporan ini dibuat dalam suatu sistematika yang terdiri dalam beberapa BAB, yaitu:

BAB I PENDAHULUAN

Pada BAB I PENDAHULUAN berisi mengenai latar berlakang masalah, identifikasi masalah, tujuan dan manfaat, ruang lingkup masalah, serta sistematika penulisan dari laporan Tugas Akhir.

BAB II LANDASAN TEORI

Pada BAB LANDASAN TEORI berisi *State of The Art (SOTA)*, tinjauan pustaka, landasan teori, kajian pendekatan yang diusulkan, dan karakteristik data yang digunakan pada laporan Tugas Akhir.

BAB III METODOLOGI PENELITIAN

Pada BAB METODOLOGI PENELITIAN berisi diagram alur metode penelitian, indikator penelitian, tahap-tahap penelitian laporan Tugas Akhir.

BAB IV PEMBAHASAN DAN HASIL

Pada BAB PEMBAHASAN & HASIL berisi hasil dan pembahasan pada laporan Tugas Akhir.

BAB V PENUTUP

Pada BAB PENUTUP berisi kesimpulan dan saran pada laporan Tugas Akhir.

DAFTAR PUSTAKA

Berisikan referensi yang menjadi dasar teori dan pikiran peneliti dalam menyusun Tugas Akhir.

LAMPIRAN-LAMPIRAN

Berisikan bukti sejumlah lampiran yang dibutuhkan sebelum melaksanakan sidang Tugas Akhir.

BAB II

TINJAUAN PUSTAKA

2.1 State of The Art (SoTA)

Teknologi informasi menjadi tantangan berat bagi pengguna teknologi informasi dengan perkembangnya dari tahun ke ketahun. Akibatnya pengguna teknologi harus lebih memahami dan mengikuti arus perkembangan tersebut. Seiring dengan berkembangnya teknologi, banyak yang mempengaruhi teknologi infomasi yang ada. Baik dibidang bisnis, hiburan, pendidikan, industri pelayanan dan banyak lagi. Salah satu contoh yang dapat diambil yaitu pada bidang industri musik. Industri musik selalu berubah dari tahun ke tahun dan menjadi salah satu dampak yang sangat berpengaruh dan juga sangat populer dikalangan pengguna teknologi saat ini. Layanan musik digital merupakan segmen penting dalam industri musik, hal ini dikarenakan memiliki manfaat seperti *streaming online* termasuk ukuran audiens, paparan yang lebih baik, konten yang kaya, berbagai *platform*, dan interaksi yang lebih baik dengan target pasar.(Noviani et al., 2020) Lagu memiliki kemampuan untuk mengekspresikan emosi manusia dan memberikan kenikmatan bagi pendengarnya. Akibatnya, sangat penting bahwa sebuah lagu tidak diputar lagi. Setelah itu, sistem rekomendasi dikembangkan untuk menjawab kebutuhan masyarakat.(Fawzan et al., 2021) *Spotify* merupakan salah satu layanan musik digital yang banyak digemari penggunanya, hal ini dikarenakan *spotify* kaya akan daftar musik yang banyak dan dapat didengarkan pengguna. Penelitian ini bertujuan untuk membuat sistem yang dapat menyajikan rekomendasi musik sesuai dengan preferensi user sehingga tingkat kenyamanan user akan meningkat.

Tim peneliti telah mengkaji sejumlah referensi mengenai perkembangan teknologi mengenai rekomendasi sistem, serta pemodelan untuk rekomendasi sistem dengan pendekatan *demographic filtering* dan *content based filtering*. Selain itu, terkait dengan kajian literatur terkait dalam melakukan rekomendasi sistem, faktor-faktor apa saja yang mempengaruhi dan model rekomendasi yang digunakan lebih lanjut. Detail dari tiap referensi lainnya ditunjukkan pada tabel. (Pereira & Varma, 2019) - (Luo, 2018)

2.2 Tinjauan Pustaka

Tabel II- 1 Tinjauan Pustaka

No.	Area Penelitian	Tahun	Karakteristik Data	Metode	Model	Hasil Penelitian
1.	Sistem rekomendasi perencanaan keuangan. (Pereira & Varma, 2019)	2019	Data kuesioner	<i>Recommendation System</i>	<i>Content-based Filtering</i>	Sistem ini berjalan di perangkat Android. Dengan penambahan sistem ini diharapkan dapat mempermudah masyarakat untuk mengunjungi apotek yang lebih dekat dengan rumahnya, memiliki fasilitas yang lebih banyak, dan memberikan pelayanan kesehatan.
2.	Mengganti model <i>Hybrid</i> untuk rekomendasi yang dipersonalisasi dengan menggabungkan pengguna	2022	Data <i>GoodReads</i> .	<i>Recommendation System</i>	<i>Content-Based Filtering</i> .	Hasil menunjukkan bahwa model hibrida mengungguli semua model tradisional lainnya dalam hal prediksi dan kualitas rekomendasi.

	<i>Demographic.</i> (K & C, 2022)					
3.	Kepercayaan User pada Sistem Rekomendasi: Sebuah Perbandingan dari <i>Content-Based</i> , <i>Collaborative</i> dan <i>Demographic Filtering</i> . (Liao et al., 2022)	2022	Data Rating	<i>Recommendation System</i>	<i>Content-Based Filtering</i>	<p>Studi saat ini juga membantu memperluas literatur saat ini tentang bagaimana pengguna sendiri memahami proses rekomendasi dan sejauh mana mereka menetapkan tanggung jawab untuk diri mereka sendiri vs. sistem, dan bagaimana atribusi ini pada gilirannya mempengaruhi kepercayaan mereka pada rekomendasi. Dengan berfokus pada faktor psikologis seperti itu dan evaluasi subjektif pengguna terhadap sistem pemberi rekomendasi, kami menawarkan wawasan praktis untuk membangun sistem tepercaya yang lebih berpusat pada manusia.</p>
4.	Meningkatkan Kuesioner yang Dipersonalisasi Dengan Pengurangan	2021	Data MovieLens	<i>Recommendation System</i>	<i>Content-Based Filtering</i>	Dalam pekerjaan ini kami meningkatkan pendekatan kuesioner yang dipersonalisasi untuk memecahkan masalah pengguna yang dingin menggunakan strategi pengurangan ketidakpastian. Hasil eksperimen pada

	Redundansi untuk Mengatasi Masalah Pengguna yang Dingin (Kabiru & Muhammad, 2021)					dataset MovieLens menandakan peningkatan RMSE, Precision, Recall, F1 dan NDCG masing-masing sebesar 0.200, 0.227, 0.261, 0.174 dan 0.249 pada dataset 1M dan 0.168, 0.159, 0.205, 0.076 dan 0.213 dalam dataset 10M masing-masing pada pekerjaan kuesioner yang dipersonalisasi dari Pozo (2018).
5.	Rekomendasi Sistem menggunakan Hybrid Filtering (Toradmall et al., 2021)	2021	Dataset data yang dihasilkan oleh pengguna.	<i>Recommendation System</i>	<i>Content-based filtering</i>	Sistem yang diusulkan akan membantu mengatasi kelemahan yang dihadapi oleh masing-masing metode, sehingga meningkatkan akurasi sistem.
6.	Kajian Sistem Rekomendasi Pada Keanekaragaman Podcast (Huda, 2021)	2021	Dataset <i>podcast</i>	<i>Recommendation System</i>	<i>Content-Based filtering</i>	Hasil penelitian menyimpulkan bahwa tingkat keragaman podcast dan perilaku pendengarnya menciptakan peluang eksplorasi karena sedikitnya studi terkait untuk menjawab tantangan yang terus berkembang.

7.	Sistem Rekomendasi untuk Item yang dapat disesuaikan (Sharma et al., 2021)	2020	Dataset <i>MovieLens</i>	<i>Recommendation System</i>	<i>Content-based filtering</i>	Nilai RMSE dari model yang diusulkan dibandingkan dengan teknik rekomendasi umum lainnya yang ada. Nilai RMSE dari model yang diusulkan dibandingkan dengan teknik rekomendasi umum lainnya seperti Content-based filtering memiliki nilai RMSE 0.644 dan Demographic filtering 0.610.
8.	Tinjauan Komprehensif Pendekatan dan Tantangan Sistem Rekomendasi (Narke & Nasreen, 2020)	2020	Data yang diambil dari database Internet/Jaringan Sosial, Media	<i>Recommendation System</i>	<i>Content-based filtering</i>	Tantangan masing-masing seperti cold start dan skala pada dataset yang lebih besar dibahas dan telah ditemukan bahwa model hybrid efektif dalam memecahkan masalah ini. Lingkup pekerjaan di masa depan termasuk menemukan lebih banyak model pelengkap yang dapat membantu sistem ini dalam mencapai skala yang lebih besar dan akurasi yang lebih baik, ada juga kelangkaan metrik perbandingan model ini. Hanya menggunakan metrik akurasi sederhana tidak menunjukkan seberapa baik model

						telah mengidentifikasi beberapa hubungan semantik lokal.
9.	Sistem Rekomendasi Film Menggunakan Pembelajaran Semi Supervised (Roy et al., 2019)	2019	Dataset film yang dirilis pada atau sebelum Juli 2017 dari Movies Dataset.	<i>Recommendation System</i>	<i>Content-Based Filtering.</i>	Penelitian ini menyajikan sistem rekomendasi untuk film menggunakan pendekatan yang berbeda dalam paradigma pemrograman Python. Sistem rekomendasi yang baik harus memberikan rekomendasi yang positif dan bermanfaat dan juga memberikan rekomendasi yang berbeda dari apa yang sudah disediakan oleh algoritma yang ada.
10.	Kepercayaan User pada Sistem Rekomendasi: Sebuah Perbandingan dari <i>Content-Based</i> , <i>Collaborative</i> dan <i>Demographic</i>	2018	Data Rating	<i>Recommendation System</i>	<i>Content-Based Filtering.</i>	Studi saat ini juga membantu memperluas literatur saat ini tentang bagaimana pengguna sendiri memahami proses rekomendasi dan sejauh mana mereka menetapkan tanggung jawab untuk diri mereka sendiri vs. sistem, dan bagaimana atribusi ini pada gilirannya mempengaruhi kepercayaan mereka pada rekomendasi. Dengan berfokus pada faktor psikologis seperti itu dan evaluasi subjektif pengguna

	<i>Filtering.</i> (Luo, 2018)				terhadap sistem pemberi rekomendasi, kami menawarkan wawasan praktis untuk membangun sistem tepercaya yang lebih berpusat pada manusia.
--	-------------------------------	--	--	--	---

2.3 Landasan Teori

Landasan teori adalah konsep dengan rapi pernyataan dan sistematis memiliki variabel dalam penelitian karena landasan teori menjadi landasan yang kuat dalam penelitian yang akan dilakukan.

2.3.1 Spotify

Spotify adalah layanan *streaming musik* yang menyediakan akses cepat ke lebih dari 8 juta lagu. *Streaming* dilakukan dengan menggabungkan akses *client-server* dan protokol *peer-to-peer*. (Kreitz & Niemelä, 2010) sebagai *platform* dengan jumlah pengguna yang banyak, memerlukan penelitian tambahan terhadap layanan *streaming* musik yang disediakan agar dapat memberikan pengalaman pengguna yang lebih baik dan meningkatkan pendapatan dibandingkan *platform streaming* lainnya melalui data mining.(Zaidah et al., 2021)

2.3.1.1 Track Name

Track name merupakan judul atau nama lagu *track*.(Maringka et al., 2021)

2.3.1.2 Album name

Album name adalah nama album lagu.(SpotifyAB, 2022)

2.3.1.3 Artist name

Artist name merupakan nama pencipta dan penyanyi lagu. (Maringka et al., 2021)

2.3.1.4 Duration

Duration adalah durasi lagu dalam hitungan milidetik milisekon.(Maringka et al., 2021)

2.3.1.5 Popularity

Popularity merupakan popularitas trek. Nilainya antara 0 dan 100, dengan 100 menjadi yang paling populer. Popularitas dihitung dengan algoritma dan sebagian besar didasarkan pada jumlah total pemutaran yang dimiliki trek dan seberapa baru pemutaran tersebut. (Stetler, 2022)

2.3.1.6 *Genres*

Genres merupakan daftar genre yang dikaitkan dengan artis.(SpotifyAB, 2022)

2.3.1.7 *Danceability*

Danceability menggambarkan seberapa cocok trek untuk menari berdasarkan kombinasi elemen musik termasuk tempo, stabilitas ritme, kekuatan ketukan, dan keteraturan keseluruhan. Nilai 0,0 paling tidak dapat menari dan 1,0 paling dapat menari. (Luo, 2018)

2.3.1.8 *Energy*

Energy adalah ukuran dari 0,0 hingga 1,0 dan mewakili ukuran persepsi intensitas dan aktivitas. Biasanya, trek energik terasa cepat, keras, dan berisik. Misalnya, death metal memiliki energi yang tinggi, sementara pendahuluan Bach mendapat skor rendah pada skala. Fitur persepsi yang berkontribusi pada atribut ini termasuk rentang dinamis, kenyaringan yang dirasakan, timbre, tingkat onset, dan entropi umum. (Luo, 2018)

2.3.1.9 *Key*

Key adalah kunci trek tersebut. Integer memetakan ke pitch menggunakan notasi Kelas *Pitch* standar. Misalnya. 0 = C, 1 = C#/Db, 2 = D, dan seterusnya.(Stetler, 2022)

2.3.1.10 *Loudness*

Loudness adalah kenyaringan keseluruhan trek dalam desibel (dB). Nilai kenyaringan dirata-ratakan di seluruh trek dan berguna untuk membandingkan kenyaringan relatif trek. Kenyaringan adalah kualitas suara yang merupakan korelasi psikologis utama dari kekuatan fisik (amplitudo). Nilai tipikal berkisar antara -60 dan 0 db.(Stetler, 2022)

2.3.1.11 *Mode*

Mode menunjukkan modalitas (major atau minor) dari sebuah trek, jenis tangga nada dari mana konten melodinya berasal. Major diwakili oleh 1 dan minor adalah 0. (Stetler, 2022)

2.3.1.12 *Speechiness*

Speechiness mendeteksi keberadaan kata-kata yang diucapkan di trek. Semakin eksklusif pidato seperti rekaman (misalnya *talk show*, buku audio, puisi), semakin mendekati 1,0

nilai atributnya. Nilai di atas 0,66 menggambarkan trek yang mungkin seluruhnya terbuat dari kata-kata yang diucapkan. Nilai antara 0,33 dan 0,66 menggambarkan trek yang mungkin berisi musik dan ucapan, baik dalam bagian atau berlapis, termasuk kasus seperti musik rap. Nilai di bawah 0,33 kemungkinan besar mewakili musik dan trek non-suara lainnya.(Luo, 2018)

2.3.1.13 *Acousticness*

Acousticness merupakan ukuran kepercayaan dari 0,0 hingga 1,0 apakah trek akustik. 1,0 mewakili kepercayaan diri yang tinggi bahwa trek tersebut akustik.(Luo, 2018)

2.3.1.14 *Instrumentaness*

Instrumentaness adalah memprediksi apakah trek tidak berisi vokal. Suara "Ooh" dan "aah" diperlakukan sebagai instrumental dalam konteks ini. *Rap* atau trek kata yang diucapkan jelas "vokal". Semakin dekat nilai instrumental menjadi 1,0, semakin besar kemungkinan trek tersebut tidak berisi konten vokal. Nilai di atas 0,5 dimaksudkan untuk mewakili trek instrumental, tetapi kepercayaan lebih tinggi saat nilainya mendekati 1,0. (Stetler, 2022)

2.3.1.15 *Liveness*

Liveness untuk mendeteksi kehadiran penonton dalam rekaman. Nilai keaktifan yang lebih tinggi menunjukkan peningkatan kemungkinan bahwa trek ditampilkan secara langsung. Nilai di atas 0,8 memberikan kemungkinan yang kuat bahwa trek itu hidup.(Stetler, 2022)

2.3.1.16 *Tempo*

Perkiraan *tempo* keseluruhan trek dalam ketukan per menit (*BPM*). Dalam terminologi musik, *tempo* adalah kecepatan atau kecepatan dari bagian tertentu dan berasal langsung dari durasi ketukan rata-rata.(Luo, 2018)

2.3.1.17 Valance

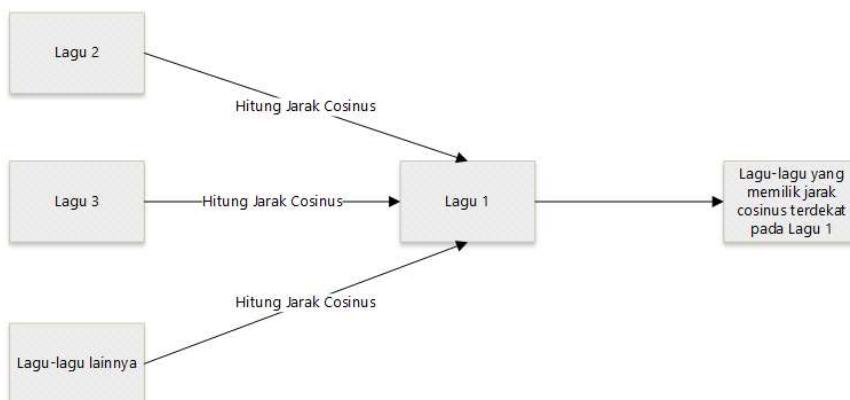
Ukuran dari 0,0 hingga 1,0 yang menggambarkan kepositifan musik yang disampaikan oleh sebuah lagu. trek dengan valensi tinggi terdengar lebih *positif* (misalnya bahagia, ceria, *euforia*), sedangkan trek dengan valensi rendah terdengar lebih negatif (misalnya sedih, tertekan, marah).(Stetler, 2022)

2.3.2 Recommendation System (Sistem Rekomendasi)

Sistem rekomendasi adalah sistem yang digunakan untuk mendapatkan sesuatu yang diinginkan.(Softscients, 2022).

2.3.3 Content-Based Filtering

Metode *content-based filtering* merupakan rekomendasi berdasarkan kemiripan konten atau item (Girsang, 2017) dan merekomendasikan konten atau item yang mirip tersebut. Sistem rekomendasi pada metode ini menghubungkan item atau konten (Adhikari, 2019) berdasarkan properti trek. Rekomendasi berbasis konten juga mampu merekomendasikan konten atau item baru yang belum pernah dinilai oleh pengguna mana pun sebelumnya karena rekomendasi ini dibuat secara eksklusif berdasarkan fitur-fiturnya, dan bukan berdasarkan peringkat pengguna lain. (Liu & Callvik, 2017) Konten atau item yang digunakan pada proses pengolahan data ini adalah *Genres*, yaitu mencari kemiripan berdasarkan *genre* yang ditentukan. *Genre index 0* pada data *top songs – global* sebagai kriteria atau parameter. *Genre* lagu lainnya akan dihitung jarak *cosinus* terdekatnya dengan lagu *index 0*. Setelah dihitung berdasarkan jarak cosinusnya maka lagu-lagu yang muncul akan menjadi lagu-lagu yang akan direkomendasikan.



Gambar II- 1 Alur Content-based filtering

2.3.3.1 Cosine Distance

Cosine distance adalah ukuran kesamaan antara dua vektor berdasarkan sudut kosinus di antara mereka. Penelitian ini mengusulkan sistem pendekripsi kemiripan *genre* dengan cara mengelompokkan dan menghitung sudut *cosinus* antar *genre* yang diperiksa. (Usino et al., 2019) *Cosine distance* digunakan untuk melakukan perhitungan kesamaan dari *genre*. Semakin besar nilai cosinus (maksimal 1) yang dihasilkan oleh *genre* yang dibandingkan, maka fitur yang dimiliki keduanya semakin mirip sehingga dapat dikelompokkan ke dalam label kelas yang sama. (Zuhraini et al., 2022) Pengukuran jarak dilakukan dengan menggunakan metode jarak cosine distance. Jarak *cosine distance* didefinisikan dengan rumus seperti berikut: (Wang et al., 2020)

$$d(x, y) = 1 - \frac{x \cdot y}{\|x\| \|y\|}$$

Dimana:

x, y = nilai rata-rata dari vektor x dan y.

2.4 Kajian Pendekatan yang diusulkan

Metode yang digunakan dalam penelitian ini adalah metode *Content-Based Filtering*. Data yang dihimpun berkenaan dengan data 50 lagu teratas- global. Tujuan penelitian ini ialah membuat *framewok* sederhana sebuah sistem rekomendasi yang berfungsi untuk menyarankan sebuah lagu berdasarkan properti trek yang ada.

2.5 Karakteristik Data yang digunakan

Dataset yang digunakan adalah data yang didapatkan dari mengkonvert data *playlist* ke dalam bentuk *format csv*. Adapun karakteristik data yang digunakan sebagai berikut: (Luo, 2018; SpotifyAB, 2022; Stetler, 2022)

Tabel II- 2 Karakteristik Data

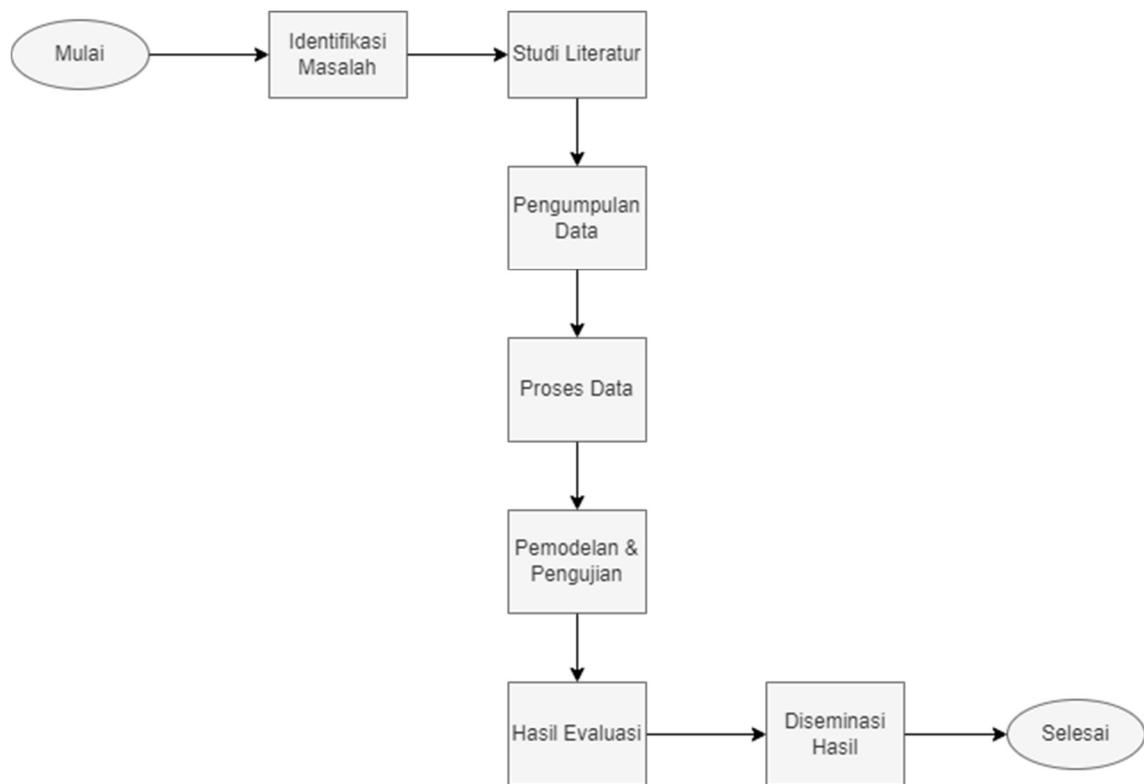
No	Atribut	Tipe Data
1.	<i>Track name</i>	<i>String</i>
2.	<i>Album name</i>	<i>String</i>
3.	<i>Artist name</i>	<i>String</i>

4.	<i>Duration</i>	<i>Int</i>
5.	<i>Popularity</i>	<i>Int</i>
6.	<i>Genres</i>	<i>String</i>
7.	<i>Danceability</i>	<i>Float</i>
8.	<i>Energy</i>	<i>Float</i>
9.	<i>Key</i>	<i>Int</i>
10.	<i>Loudness</i>	<i>Float</i>
11	<i>Mode</i>	<i>Int</i>
12.	<i>Speechiness</i>	<i>Float</i>
13.	<i>Acousticness</i>	<i>Float</i>
14.	<i>Instrumentaness</i>	<i>Float</i>
15	<i>Liveness</i>	<i>Float</i>
16.	<i>Tempo</i>	<i>Float</i>
17.	<i>Valance</i>	<i>Float</i>

BAB III

METODOLOGI PENELITIAN

3.1 Diagram Alur Metodologi Penelitian



Gambar III- 1 Diagram Alur Metodologi Penelitian

3.2 Indikator Capaian

Berdasarkan diagram alur metode penelitian diatas, terdapat indikator capaian sebagai berikut:

Tabel III- 1 Indikator Capaian

No.	Tahapan		Indikator capaian
1.	Identifikasi dan perumusan masalah	→	1. Penentuan Rumusan Masalah.

2.	Studi literatur	→	2. <i>Concept Map</i> Sistem Rekomendasi lagu terhadap properti trek yang ada berdasarkan data dan metode <i>Content-Based Filtering</i> .
3.	Pengumpulan data	→	3. Data mentah yang di <i>convert</i> dari <i>playlist platform</i> musik <i>spotify</i> .
3.	Proses Data	→	4. Proses data yang sudah siap untuk pemodelan dengan tahapan analisa, pembersihan, penanganan nilai yang hilang dan transformasi.
4.	Pemodelan & Pengujian	→	5. <i>Content-Based Filtering</i> merupakan pemodelan yang digunakan untuk merekomendasikan lagu berdasarkan properti trek. Data diuji menggunakan <i>Jupyter Notebook</i> dalam mengolah data menggunakan pemodelan.
5.	Evaluasi	→	6. Hasil penelitian dan performansi model.
6.	Diseminasi hasil	→	7. Laporan Tugas Akhir dan <i>draft</i> jurnal Tugas Akhir.

3.3 Tahapan-Tahapan Metode Penelitian

Berikut adalah tahapan-tahapan metode penelitian:

3.3.1 Identifikasi Masalah

Identifikasi masalah merupakan langkah pertama dalam tahapan metode penelitian yang berfungsi untuk mengidentifikasi masalah. Pada tahap ini peneliti melakukan penelitian dengan cara memunculkan masalah (*problem*) yang ada, kemudian melakukan identifikasi sumber permasalahan (*root cause*) sehingga menciptakan permasalahan (*problem statement*) yang menjelaskan permasalahan yang sudah diidentifikasi. Adapun Permasalahan yang terdapat pada penelitian ini adalah bagaimana cara merekomendasikan lagu berdasarkan properti trek.

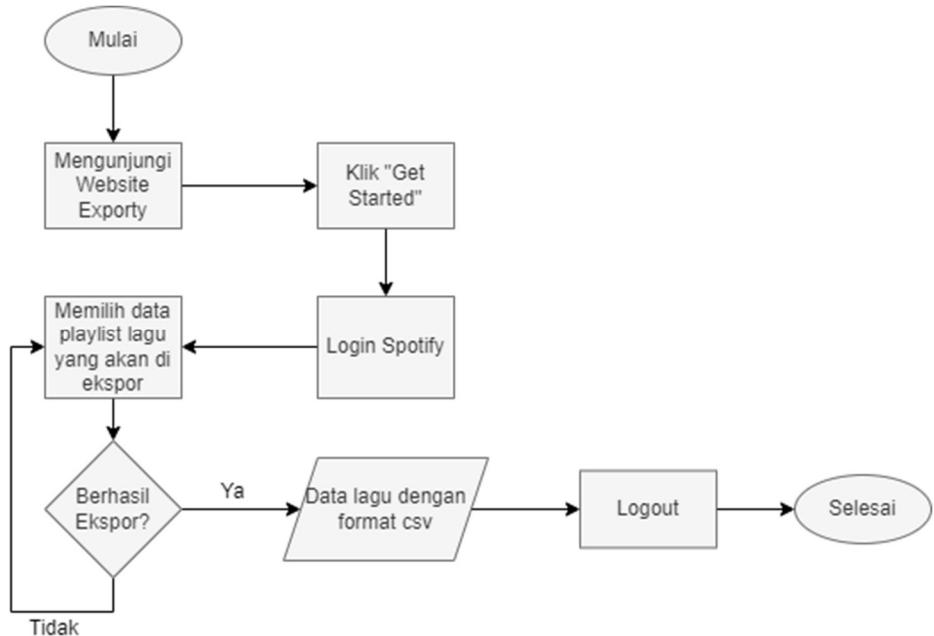
3.3.2 Studi Literatur

Setelah dilakukan tahap indentifikasi masalah langkah selanjutnya dalam tahapan metode penelitian adalah studi literatur. Pada tahap ini peneliti melakukan pencarian informasi yang relevan dengan masalah yang akan diteliti dan mengkaji beberapa teori dasar yang relevan dengan permasalahan yang akan diteliti kemudian memperdalam pengetahuan peneliti tentang permasalahan yang akan diteliti dan mengkaji hasil-hasil penelitian terdahulu yang berkaitan dengan penelitian yang akan diteliti dan mendapatkan informasi tentang aspek-aspek mana dari suatu permasalahan yang sudah pernah diteliti untuk menghindari agar tidak terjadi *plagiarisme* penelitian.

3.3.3 Pengumpulan Data

Pengumpulan data adalah tahap setelah studi literatur. Tujuan pengumpulan data adalah untuk memperoleh data yang akan digunakan dalam penelitian. Peneliti melakukan pengumpulan studi literatur yang relevansi dan sesuai dengan apa yang dibutuhkan untuk menunjang penelitian untuk dapat melakukan teknik pengumpulan data studi literatur. Maka dari itu data yang digunakan pada penelitian ini adalah data yang di *convert* dari website <https://exportify.net/#playlists> kedalam bentuk *format csv*. Website tersebut

berfungsi untuk mengekspor daftar putar *spotify* menggunakan *Web API*. Alur dari pengumpulan data dapat dilihat dari *flowchart* berikut:



Gambar III- 2 Flowchart pengumpulan data

Penjelasan *flowchart*, dimulai dengan cara mengunjungi website *Exportify* kemudian akan muncul tampilan awal dari website. Klik “*Get Started*”, kemudian akan dialihkan ke halaman *login* dan melakukan proses *login*. Setelah masuk *login* selanjutnya akan ditampilkan daftar *spotify* yang sudah ditambahkan kedalam koleksi *playlist* pada *spotify*. Kemudian klik *export* lalu *playlist* yang berisi lagu yang ingin diambil datanya akan terekspor menjadi data yang berbentuk *file csv*. kemudian *logout* setelah selesai melakukan *export* data.

3.3.4 Proses Data

Setelah dilakukan pengumpulan data langkah selanjutnya dalam tahap metodologi penelitian adalah proses data. Pada tahap proses data atau pengolahan data dilakukan data *cleaning*. *Cleaning* data adalah prosedur untuk memastikan keakuratan, konsistensi, dan kegunaan data dalam kumpulan data. Prosedurnya adalah mendeteksi kesalahan atau kerusakan pada data dan kemudian memperbaiki atau menghapusnya jika perlu.

3.3.5 Pemodelan dan Pengujian

Tahapan selanjutnya yaitu Pemodelan dan Pengujian. Tahap ini ialah tahap menggunakan metode rekomendasi dalam pemodelan dan pengujian data. *Content-Based Filtering* merupakan metode, model atau algoritma yang akan digunakan pada penelitian ini. Pemodelan ada tahap yang dilakukan sebelum pengujian. Dalam pemodelan dipertimbangkan dulu model yang akan dipakai dalam mengelola data. Kemudian langkah selanjutnya setelah pemodelan adalah langkah Pengujian menggunakan model *Content-Based Filtering*.

3.3.6 Hasil Evaluasi

Tahap hasil evaluasi adalah tahap performansi dari pemodelan dan pengujian. Hasil pengujian data akan diimplementasi pada tahap hasil evaluasi.

3.3.7 Diseminasi Hasil

Tahap deseminasi hasil adalah tahap untuk artikel yang diterbitkan dalam laporan penelitian dan *draft* jurnal.

BAB IV

PEMBAHASAN DAN HASIL

4.1 Pembahasan

Penelitian ini dilakukan dengan mengambil data selama 1 minggu terakhir dari *playlist* 50 Lagu Teratas - Global (*50 Top Songs - Global*) sebanyak 50 data lagu. Untuk hasil data mentah atau data yang belum dilakukan proses data yaitu data yang ada pada *spotify*, dapat dilihat sebagai berikut:

The screenshot shows a Spotify playlist titled "Lagu Teratas - Global". The interface includes a sidebar for "Top Songs Global" and a main area for the "Weekly Music Charts". The playlist details are as follows:

#	JUDUL	PUTERAN	ALBUM	DURASI
1	Quevedo: Bzrp Music Sessions, Vol. 52	49.200.772	Quevedo: Bzrp Music Sessions, Vo...	3:18
2	As It Was	39.156.789	Harry's House	2:47
3	Me Porto Bonito	36.028.246	Un Verano Sin Ti	2:58
4	Titi Me Preguntó	32.101.875	Un Verano Sin Ti	4:03
5	Running Up That Hill (A Deal With God)	30.032.752	Hounds Of Love	4:58
6	DESPECHÁ	28.107.868	DESPECHÁ	2:37
7	Efecto	26.359.490	Un Verano Sin Ti	3:33
8	Glimpse of Us	25.333.978	Glimpse of Us	3:53
9	Ojitos Lindos	23.486.675	Un Verano Sin Ti	4:18
10	Bad Habit	21.288.974	Gemini Rights	3:52
11	Moscow Mule	20.845.139	Un Verano Sin Ti	4:05
12	Heat Waves	20.321.676	Dreamland (+ Bonus Levels)	3:58
13	La Bachata	20.180.692	La Bachata	2:42
14	PROVENZA	19.745.401	PROVENZA	3:30
15	Bad Decisions (with BTS & Snoop Dogg)	19.500.644	Bad Decisions (with BTS & Snoop ...	2:52
16	I Ain't Worried	19.361.468	I Ain't Worried (Music From The M...	2:28

17		Left and Right (Feat. Jung Kook of BTS)	18.378.990	Left and Right (Feat. Jung Kook of ...	2:54
18		Doja	17.262.833	Doja	1:37
19		BREAK MY SOUL	16.262.031	RENAISSANCE	4:38
20		About Damn Time	15.991.437	Special	3:11
21		STAYING ALIVE (feat. Drake & Lil Baby)	15.991.390	STAYING ALIVE (feat. Drake, Lil ...	2:58
22		STAY (with Justin Bieber)	15.847.791	F*CK LOVE 3: OVER YOU	2:21
23		Tarot	15.746.632	Un Verano Sin Ti	3:57
24		Cold Heart - PNAU Remix	15.351.197	The Lockdown Sessions	3:22
25		Te Felicito	15.230.719	Te Felicito	2:52
26		SNAP	15.219.493	SNAP PACK	2:59
27		Late Night Talking	14.395.673	Harry's House	2:57
28		Party	14.346.290	Un Verano Sin Ti	3:47
29		I Like You (A Happier Song) (with Doja Cat)	14.133.871	Twelve Carat Toothache	3:12
30		Bam Bam (feat. Ed Sheeran)	14.016.418	Familia	3:26
31		Jimmy Cooks (feat. 21 Savage)	13.188.576	Honestly, Nevermind	3:38
32		Under The Influence	13.145.783	Indigo (Extended)	3:04
33		Sunroof	13.141.161	Sunroof	2:43
34		Sweater Weather	12.617.763	I Love You.	4:00
35		Blinding Lights	12.478.897	After Hours	3:20
36		Ferrari	12.308.447	Ferrari	3:06
37		Ghost	12.221.064	Justice	2:33
38		La Corriente	12.129.312	Un Verano Sin Ti	3:18
39		One Kiss (with Dua Lipa)	11.931.800	One Kiss (with Dua Lipa)	3:34
40		Vegas (From the Original Motion Picture Soundtrack ...)	11.912.237	Vegas (From the Original Motion P...	3:02
41		Die For You	11.891.349	Starboy	4:20
42		Until I Found You	11.831.698	Until I Found You	2:57
43		MIDDLE OF THE NIGHT	11.830.078	MIDDLE OF THE NIGHT	3:04
44		First Class	11.457.282	Come Home The Kids Miss You	2:53
45		Another Love	11.450.164	Long Way Down (Deluxe)	4:04
46		Una Noche en Medellín	11.162.073	Una Noche en Medellín	2:33
47		Kesariya (From "Brahmastra")	10.934.301	Kesariya (From "Brahmastra")	4:28
48		Despúes de la Playa	10.923.650	Un Verano Sin Ti	3:50
49		Dandelions	10.619.440	Safe Haven	3:53
50		Shivers	10.503.440	=	3:27

Gambar IV- 1 Playlist spotify lagu teratas -global

Kemudian data *diconvert* dalam *format file csv*.

Tabel IV- 1Playlist spotify berbentuk file csv

Spotify ID	Artist IDs	Track Name	Album Name	Artist Name(s)	Release Date	Duration
2tTmW7RDtMQtBk7m2rYeSw	"716NhGYqD1jl2wl1Qkgq36,52iwsT98xCoGgiG					
4Dvkj6JhhA12EX05fT7y2e	"6KImCVD70vtloJWnq6nGn3"	"As It Was"	"Harry's			
6Sq7ltF9Qa7SNFBsV5Cogx	"4q3ewBCX7sLwd24euuV69X,37230BxxYs9ksS7Olk					
1IHWI5LamUGEuP4ozKQSXZ	"4q3ewBCX7sLwd24euuV69X"	"Tití Me Preguntó"				
75FEaRjZTKLhTrFGsfMUXR	"1aSxMhuvixZ8h9dk9jIDwL"	"Running Up That Hill"				
5ildQOEKmJuWGI2vRkFdYc	"7ltDVBr6mKbRvohxheJ9h1"	"DESPECHÁ"	"DESPE			
5Eax0qFko2dh7RI2lYs3bx	"4q3ewBCX7sLwd24euuV69X"	"Efecto"	"Un Verano			
6xGruZOHLs39ZbVccQTuPZ	"3MZsBdqDrRTJihTHQrO6Dq"	"Glimpse of Us"	"Glimpse of Us"			
3k3NWokhRRkEPhCzPmV8TW	"4q3ewBCX7sLwd24euuV69X,5n9bMYfz9qss2"					

Berikut untuk data dalam bentuk tabel:

Tabel IV-2 Data berbentuk tabel

<i>Track Name</i>	<i>Album Name</i>	<i>Artist Name(s)</i>	<i>Popularit y</i>	<i>Genres</i>	<i>Danceabilit y</i>	<i>Energy</i>	<i>Valence</i>	<i>Tempo</i>
<i>Quevedo: Bzrp Music Sessions, Vol. 52</i>	<i>Quevedo: Bzrp Music Sessions, Vol. 52</i>	<i>Bizarrap, Quevedo</i>	<i>100</i>	<i>argentine hip hop, pop venezolano, trap argentino, rap canario, trap latino</i>	<i>0,43125</i>	<i>0,543056</i>	<i>00.55</i>	<i>128.033</i>
<i>As It Was</i>	<i>Harry's House</i>	<i>Harry Styles</i>	<i>93</i>	<i>pop</i>	<i>00.52</i>	<i>0,507639</i>	<i>0,459722</i>	<i>173.93</i>
<i>Me Porto Bonito</i>	<i>Un Verano Sin Ti</i>	<i>Bad Bunny, Chencho Corleone</i>	<i>99</i>	<i>latin, reggaeton, trap latino</i>	<i>0,632639</i>	<i>0,494444</i>	<i>0,295139</i>	<i>92.005</i>
<i>Tití Me Preguntó</i>	<i>Un Verano Sin Ti</i>	<i>Bad Bunny</i>	<i>98</i>	<i>latin, reggaeton, trap latino</i>	<i>0,045139</i>	<i>0,496528</i>	<i>0,129861</i>	<i>106.672</i>
<i>Running Up That Hill (A Deal with God)</i>	<i>Hounds of Love</i>	<i>Kate Bush</i>	<i>95</i>	<i>art pop, art rock, baroque pop, new wave pop, permanent wave, piano rock, singer-songwriter</i>	<i>0,436806</i>	<i>0,379861</i>	<i>0,136806</i>	<i>108.375</i>

<i>DESPECHÁ</i>	<i>DESPECHÁ</i>	<i>ROSALÍA</i>	93	<i>r&b en espanol</i>	0,638194	0,43263 9	0,53819 4	130.037
<i>Efecto</i>	<i>Un Verano Sin Ti</i>	<i>Bad Bunny</i>	96	<i>latin,reggaeton,trap latino</i>	0,55625	0,32986 1	0,1625	98.047
<i>Glimpse of Us</i>	<i>Glimpse of Us</i>	<i>Jojí</i>	97	<i>viral pop</i>	00.44	0,22013 9	0,18611 1	169.914
<i>Ojitos Lindos</i>	<i>Un Verano Sin Ti</i>	<i>Bad Bunny,Bomba Estéreo</i>	96	<i>latin,reggaeton,trap latino,cumbia,latin alternative,pop electronico,tropical alternativo</i>	0,449306	0,47638 9	0,18611 1	79.928
<i>Bad Habit</i>	<i>Gemini Rights</i>	<i>Steve Lacy</i>	85	<i>afrofuturism</i>	0,476389	0,34305 6	00.07	168.946
<i>Moscow Mule</i>	<i>Un Verano Sin Ti</i>	<i>Bad Bunny</i>	95	<i>latin,reggaeton,trap latino</i>	0,558333	0,46805 6	0,20277 8	99.968
<i>Heat Waves</i>	<i>Dreamland (+ Bonus Levels)</i>	<i>Glass Animals</i>	90	<i>gauze pop,indietronica,shiver pop</i>	0,528472	0,36458 3	0,36875	80.87
<i>La Bachata</i>	<i>La Bachata</i>	<i>Manuel Turizo</i>	92	<i>colombian pop,latin,latin pop,reggaeton,reggaeton colombiano,trap latino</i>	0,579861	0,47152 8	0,05902 8	124.98

<i>PROVENZ A</i>	<i>PROVENZA</i>	<i>KAROL G</i>	<i>94</i>	<i>latin,reggaeton,reggaeton colombiano</i>	<i>0,060417</i>	<i>0,35833 3</i>	<i>00.53</i>	<i>111.005</i>
<i>Bad Decisions (with BTS & Snoop Dogg)</i>	<i>Bad Decisions (with BTS & Snoop Dogg)</i>	<i>benny blanco,BTS,Snoo p Dogg</i>	<i>88</i>	<i>electropop,pop,pop rap,k- pop,k-pop boy group,g funk,gangster rap,hip hop,rap,west coast rap</i>	<i>0,529167</i>	<i>0,59791 7</i>	<i>0,66319 4</i>	<i>120.048</i>
<i>I Ain't Worried</i>	<i>I Ain't Worried (Music from The Motion Picture Top Gun: Maverick)</i>	<i>OneRepublic</i>	<i>93</i>	<i>dance pop,piano rock,pop,pop rock</i>	<i>0,488889</i>	<i>0,55347 2</i>	<i>0,57291 7</i>	<i>139.994</i>
<i>Left and Right (Feat. Jung Kook of BTS)</i>	<i>Left and Right (Feat. Jung Kook of BTS)</i>	<i>Charlie Puth,Jung Kook,BTS</i>	<i>94</i>	<i>dance pop,pop,viral pop,k- pop,k-pop boy group</i>	<i>0,611806</i>	<i>0,41111 1</i>	<i>0,49930 6</i>	<i>101.058</i>
<i>Doja</i>	<i>Doja</i>	<i>Central Cee</i>	<i>91</i>	<i>drill brasileiro</i>	<i>0,632639</i>	<i>0,39791 7</i>	<i>0,675</i>	<i>140.04.0 0</i>

<i>BREAK MY SOUL</i>	<i>RENAISSANCE</i>	<i>Beyoncé</i>	82	<i>dance pop,pop,r&b</i>	0,48125	0,61597 2	0,6	114.942
<i>About Damn Time</i>	<i>Special</i>	<i>Lizzo</i>	80	<i>dance pop,escape room,minnesota hip hop,pop,trap queen</i>	0,580556	0,51597 2	0,50208 3	108.965
<i>STAYING ALIVE (feat. Drake & Lil Baby)</i>	<i>STAYING ALIVE (feat. Drake & Lil Baby)</i>	<i>DJ Khaled,Drake,Lil Baby</i>	85	<i>dance pop,hip hop,miami hip hop,pop,pop rap,rap,southern hip hop,trap,canadian hip hop,canadian pop,toronto rap,atl hip hop,atl trap</i>	0,498611	0,32013 9	0,12569 4	129.953
<i>STAY (with Justin Bieber)</i>	<i>F*CK LOVE 3: OVER YOU</i>	<i>The Kid LAROI,Justin Bieber</i>	86	<i>australian hip hop,canadian pop,pop</i>	0,410417	0,53055 6	0,33194 4	169.928
<i>Tarot</i>	<i>Un Verano Sin Ti</i>	<i>Bad Bunny,Jhay Cortez</i>	93	<i>latin,reggaeton,trap latino</i>	0,552083	0,475	0,29097 2	114.011
<i>Cold Heart - PNAU Remix</i>	<i>The Lockdown Sessions</i>	<i>Elton John,Dua Lipa,PNAU</i>	84	<i>glam rock,mellow gold,piano rock,dance pop,pop,uk pop,alternative dance,aussietronica,australi</i>	0,552083	00.08	0,64861 1	116.032

				<i>an dance,australian electropop</i>				
<i>Te Felicito</i>	<i>Te Felicito</i>	<i>Shakira,Rauw Alejandro</i>	<i>93</i>	<i>colombian pop,dance pop,latin,latin pop,pop,puerto rican pop,trap latino</i>	<i>0,482639</i>	<i>0,44166 7</i>	<i>0,39930 6</i>	<i>174.14.0 0</i>
<i>SNAP</i>	<i>SNAP PACK</i>	<i>Rosa Linn</i>	<i>72</i>		<i>0,392361</i>	<i>0,44166 7</i>	<i>0,36458 3</i>	<i>170.01.0 0</i>
<i>Late Night Talking</i>	<i>Harry's House</i>	<i>Harry Styles</i>	<i>93</i>	<i>pop</i>	<i>0,495833</i>	<i>0,50555 6</i>	<i>0,62569 4</i>	<i>114.996</i>
<i>Party</i>	<i>Un Verano Sin Ti</i>	<i>Bad Bunny,Rauw Alejandro</i>	<i>93</i>	<i>latin,reggaeton,trap latino,puerto rican pop</i>	<i>0,575694</i>	<i>0,55486 1</i>	<i>0,32708 3</i>	<i>97.013</i>
<i>I Like You (A Happier Song) (with Doja Cat)</i>	<i>Twelve Carat Toothache</i>	<i>Post Malone,Doja Cat</i>	<i>91</i>	<i>dfw rap,melodic rap,rap,dance pop,pop</i>	<i>0,509028</i>	<i>0,04652 8</i>	<i>0,32777 8</i>	<i>100.964</i>
<i>Bam Bam (feat. Ed Sheeran)</i>	<i>Familia</i>	<i>Camila Cabello,Ed Sheeran</i>	<i>84</i>	<i>dance pop,pop,post-teen pop,uk pop</i>	<i>0,525</i>	<i>0,48402 8</i>	<i>0,66388 9</i>	<i>94.996</i>

<i>Jimmy Cooks (feat. 21 Savage)</i>	<i>Honestly, Nevermind</i>	<i>Drake,21 Savage</i>	92	<i>canadian hip hop,canadian pop,hip hop,rap,toronto rap,atl hip hop,trap</i>	0,367361	0,467361	0,254167	165.921
<i>Under The Influence</i>	<i>Indigo (Extended)</i>	<i>Chris Brown</i>	88	<i>dance pop,pop,pop rap,r&b</i>	0,509028	0,047917	00.31	116.992
<i>Sunroof</i>	<i>Sunroof</i>	<i>Nicky Youre,dazy</i>	91		0,533333	0,495833	0,584722	131.443
<i>Sweater Weather</i>	<i>I Love You.</i>	<i>The Neighbourhood</i>	90	<i>modern alternative rock,modern rock,pop,shimmer pop</i>	0,425	0,560417	0,276389	124.053
<i>Blinding Lights</i>	<i>After Hours</i>	<i>The Weeknd</i>	90	<i>canadian contemporary r&b,canadian pop,pop</i>	0,356944	0,050694	0,231944	171.005
<i>Ferrari</i>	<i>Ferrari</i>	<i>James Hype,Miggy Dela Rosa</i>	90	<i>deep groove house,edm,house,pop dance,tropical house,uk dance</i>	0,588194	0,047917	0,480556	125.004
<i>Ghost</i>	<i>Justice</i>	<i>Justin Bieber</i>	89	<i>canadian pop,pop</i>	0,417361	0,514583	0,30625	153.96
<i>La Corriente</i>	<i>Un Verano Sin Ti</i>	<i>Bad Bunny,Tony Dize</i>	91	<i>latin,reggaeton,trap latino,latin hip hop</i>	0,460417	0,545833	0,402083	196.12.00

<i>One Kiss (with Dua Lipa)</i>	<i>One Kiss (with Dua Lipa)</i>	<i>Calvin Harris,Dua Lipa</i>	90	<i>dance pop,edm,electro house,house,pop,progressive house,uk dance,uk pop</i>	0,549306	0,59861 1	0,41111 1	123.994
<i>Vegas (From the Original Motion Picture Soundtrack ELVIS)</i>	<i>Vegas (From the Original Motion Picture Soundtrack ELVIS)</i>	<i>Doja Cat</i>	90	<i>dance pop,pop</i>	0,55625	0,41736 1	0,05138 9	159.969
<i>Die for You</i>	<i>Starboy</i>	<i>The Weeknd</i>	84	<i>canadian contemporary r&b,canadian pop,pop</i>	0,406944	0,36458 3	0,35277 8	133.629
<i>Until I Found You</i>	<i>Until I Found You</i>	<i>Stephen Sanchez</i>	91	<i>gen z singer-songwriter</i>	0,374306	0,35277 8	0,15763 9	101.358
<i>MIDDLE OF THE NIGHT</i>	<i>MIDDLE OF THE NIGHT</i>	<i>Elley Duhé</i>	91	<i>alt z,pop</i>	00.41	0,42430 6	0,62430 6	185.727
<i>First Class</i>	<i>Come Home the Kids Miss You</i>	<i>Jack Harlow</i>	82	<i>deep underground hip hop,kentucky hip hop,rap</i>	0,626389	0,40416 7	0,23055 6	107.005

<i>Another Love</i>	<i>Long Way Down</i>	<i>Tom Odell</i>	55	<i>chill pop, pop</i>	0,306944	0,37361 1	0,09652 8	122.758
<i>Una Noche en Medellín</i>	<i>Una Noche en Medellín</i>	<i>Cris Mj</i>	91	<i>mambo chileno, urbano chileno</i>	0,060417	0,38055 6	0,57777 8	96.018
<i>Kesariya (From Brahmastra)</i>	<i>Kesariya (From Brahmastra)</i>	<i>Pritam,Arijit Singh,Amitabh Bhattacharya</i>	87	<i>desi pop,film,indian instrumental,modern bollywood</i>	0,400694	0,39861 1	0,30277 8	94.009
<i>Después de la Playa</i>	<i>Un Verano Sin Ti</i>	<i>Bad Bunny</i>	91	<i>latin,reggaeton,trap latino</i>	0,391667	0,62708 3	0,42152 8	78.293
<i>Dandelions</i>	<i>Safe Haven</i>	<i>Ruth B.</i>	90	<i>alt z,canadian contemporary r&b,canadian pop, pop</i>	0,422917	0,48055 6	0,31527 8	116.959
<i>Shivers</i>	=	<i>Ed Sheeran</i>	89	<i>pop,uk pop</i>	0,547222	0,59652 8	0,57083 3	141.02.0 0

Proses selanjutnya dalam penelitian ini adalah proses pemodelan dan pengujian data. Penelitian ini menggunakan *content-based filtering* dalam pemodelannya. Langkah-langkah dalam pemodelan menggunakan *content-based filtering*.

a) Import Library

```
In [1]: import pandas as pd
import numpy as np
import sklearn
import nltk
from sklearn.metrics.pairwise import cosine_distances
from sklearn.feature_extraction.text import CountVectorizer
from nltk.tokenize import word_tokenize
```

Gambar IV- 2 Import Library

b) Menampilkan 5 dataset teratas.

	Spotify ID	Artist IDs	Track Name	Album Name	Artist Name(s)	Release Date	Duration (ms)	Popularity	Add
0	2tMw7RDTMQtBk7m2rYeSw	716NhGYqD1jI2wI1Qkgq36,52iwsT98xCoGgiGntTiR7K	Quevedo: Bzrp Music Sessions, Vol. 52	Bzrp Music Sessions, Vol. 52	Bizarrap, Quevedo	2022-07-06	198937	100	spotify
1	4Dvkj6JhhA12EX05fT7y2e	6KlmCV70vtloJWnq6nGn3	As It Was	Harry's House	Harry Styles	2022-05-20	167303	93	spotify
2	6Sq7lfF9Qa7SNFBsV5Cogx	4q3ewBCX7sLwd24euuV69X,37230BxxYs9ksS7OkZw3lU	Me Porto Bonito	Un Verano Sin Ti	Bunny, Chencho Corleone	2022-05-06	178567	99	spotify
3	1HWI5LamUGEuP4ozKQSXZ	4q3ewBCX7sLwd24euuV69X	Tilt Me Preguntó	Un Verano Sin Ti	Bad Bunny	2022-05-06	243716	98	spotify
4	75FEaRjZTKLhTrFGsfMUXR	1aSxMhuvixZ8h9dK9jIDwL	Running Up That Hill (A Deal With God)	Hounds Of Love	Kate Bush	1985	298933	95	spotify

5 rows × 23 columns

Gambar IV- 3 Menampilkan dataset

c) Cleaning Data

Cleaning data adalah prosedur untuk memastikan keakuratan, konsistensi, dan kegunaan data dalam kumpulan data. Prosedurnya adalah mendeteksi kesalahan atau kerusakan pada data dan kemudian memperbaiki atau menghapusnya jika perlu. Berikut adalah proses data yang telah dilakukan *cleaning* data:

- *Drop* tabel

Drop tabel yang befungsi untuk menghilangkan tabel yang tidak digunakan atau diperlukan dalam proses penelitian. Terdapat beberapa data yang tidak diperlukan dalam penelitian ini, dan di lakukan *drop* tabel. Dapat dilihat sebagai berikut:

Out[5]:													
	Track Name	Album Name	Artist Name(s)	release	duration	Popularity	Genres	Danceability	Energy	Key	Loudness	Mode	Spe...
0	Quevedo: Bzrp Music Sessions, Vol. 52	Quevedo: Bzrp Music Sessions, Vol. 52	Bizarrap, Quevedo	2022-07-06	198937	100	argentine hip hop, pop venezolano, trap argentin...	0.621	0.782	2	-5.548	1	
1	As It Was	Harry's House	Harry Styles	2022-05-20	167303	93	pop	0.520	0.731	6	-5.338	0	
2	Me Porto Bonito	Un Verano Sin Ti	Bad Bunny, Chenco Corleone	2022-05-06	178567	99	latin, reggaeton, trap latino	0.911	0.712	1	-5.105	0	
3	Titi Me Preguntó	Un Verano Sin Ti	Bad Bunny	2022-05-06	243716	98	latin, reggaeton, trap latino	0.650	0.715	5	-5.198	0	
4	Running Up That Hill (A Deal With God)	Hounds Of Love	Kate Bush	1985	298933	95	art pop, art rock, baroque pop, new wave pop, perm...	0.629	0.547	10	-13.123	0	

Gambar IV- 4 Drop tabel

- Cek Data

Mengecek apakah data tersebut memiliki *missing value* atau tidak. Setelah dilakukan pengecekan data terdapat 2 data yang memiliki *missing value*:

In [6]:	df.isnull().sum()
Out[6]:	Track Name 0 Album Name 0 Artist Name(s) 0 release 0 duration 0 Popularity 0 Genres 2 Danceability 0 Energy 0 Key 0 Loudness 0 Mode 0 Speechiness 0 Acousticness 0 Instrumentalness 0 Liveness 0 Valence 0 Tempo 0 Time Signature 0 dtype: int64

Gambar IV- 5 Missing Value

Lalu dilakukan proses mengisi nilai *NA/NaN* (*Not a Number*) menjadi 0.

In [7]:	df.fillna(0).head()
Out[7]:	
0	Quevedo: Bzrp Music Sessions, Vol. 52
1	As It Was
2	Me Porto Bonito
3	Titi Me Preguntó
4	Running Up That Hill (A Deal With God)

Gambar IV- 6 Mengisi nilai NA/NaN

Kemudian menghapus atau menghilangkan semua entri dengan nilai *NaN* atau 0 seperti berikut:

In [8]:	df = df.dropna(axis = 0) df.shape
Out[8]:	(48, 19)

Gambar IV- 7 Drop entri

d) Pemodelan sistem rekomendasi berdasarkan konten.

Pada pemodelan menggunakan *content-based filtering* memiliki beberapa tahap dalam menampilkan rekomendasi berdasarkan konten. Dapat dilihat sebagai berikut:

- Encode Data

Pada tahap ini *genres* di *encode* dengan cara mengkonversi *teks* menjadi *vector* menggunakan *CountVectorizer* dengan konten yang digunakan dalam merekomendasikan yaitu kemiripan konten index 0 pada properti trek *genres*.

Dengan langkah sebagai berikut:

content based filtering

encode data

Gambar IV- 8 Encode Data

- Mencari *similarity*.

Pada tahap ini kita menggunakan *cosine distance* dalam mencari kesamaan atau kemiripan konten dengan cara menghitung jarak *cosinus* yaitu ukuran kesamaan antara dua vektor berdasarkan sudut kosinus di antara konten. (Usino et al., 2019) Kemudian mengembalikan nilai index bilangan bulat yang akan mengurutkan indeks. Dari hasil mengurutkan tersebut maka didapat hasil rekomendasi berdasarkan *index* 0 yaitu 5 lagu teratas yang mirip dengan konten yang sama dengan *index* 0 dan dapat dilihat sebagai berikut:

```
search

In [15]: dist = cosine_distances(code, bank)
dist

Out[15]: array([[ 3.33066907e-16, 8.14304662e-01, 2.77814619e-01, 2.77814619e-01,
   3.63250049e-01, 1.00000000e+00, 2.77814619e-01, 8.68693567e-01,
   2.57218647e-01, 1.00000000e+00, 2.77814619e-01, 4.40107489e-01,
   2.74981512e-01, 5.30223284e-01, 2.38009986e-01, 3.14264657e-01,
   3.85817225e-01, 1.00000000e+00, 4.63943733e-01, 2.27951355e-01,
   1.64717345e-01, 3.56732479e-01, 2.77814619e-01, 3.24668997e-01,
   2.94785725e-01, 8.14304662e-01, 2.57218647e-01, 3.10641227e-01,
   4.38510075e-01, 1.99986938e-01, 3.88477526e-01, 3.57197835e-01,
   3.53761655e-01, 4.12112863e-01, 5.4511174e-01, 2.55930214e-01,
   3.48198443e-01, 5.45141174e-01, 5.35761655e-01, 1.00000000e+00,
   3.53761655e-01, 3.96490151e-01, 5.4511174e-01, 7.19255037e-01,
   3.96490151e-01, 2.77814619e-01, 4.57917835e-01, 5.45141174e-01]])
```



```
In [24]: rec_idx = dist.argsort()[0, 0:6]
rec_idx

Out[24]: array([ 0, 20, 29, 19, 14, 35], dtype=int64)
```

In [25]:	df.loc[rec_idx]													
Out[25]:	Track Name	Album Name	Artist Name(s)	release	duration	Popularity	Genres	Danceability	Energy	Key	Loudness	Mode	Speechiness	Acous
0	Quevedo: Bzrp Music Sessions Vol. 52	Quevedo: Bzrp Music Sessions Vol. 52	Bizarrap,Quevedo	2022-07-06	198937	100	argentine hip hop,pop venezolano,trap argentin...	0.621	0.782	2	-5.548	1	0.0440	
20	STAYING ALIVE (feat. Drake & Lil Baby)	STAYING ALIVE (feat. Drake & Lil Baby)	Khaled,Drake,Lil Baby	2022-08-05	178176	85	dance pop,hip hop,miami hip hop,pop,pop,rap,ra...	0.718	0.461	4	-7.676	0	0.0797	
29	Bam (feat. Ed Sheeran)	Familia	Camila Cabello,Ed Sheeran	2022-04-07	206070	84	dance pop,pop,post-teen pop,uk pop	0.756	0.697	8	-6.377	1	0.0401	
19	About Damn Time	Special	Lizzo	2022-07-15	191937	80	dance pop,escape room,minnesota hip hop,pop,tr...	0.836	0.743	10	-6.305	0	0.0653	
14	Bad Decisions (with BTS & Snoop Dogg)	Bad Decisions (with BTS & Snoop Dogg)	benny blanco,BTS,Snoop Dogg	2022-08-05	172714	88	electropop,pop,rap,k-pop,k-pop boy group,g...	0.762	0.861	0	-3.296	1	0.0955	
35	Ferrari	Ferrari	James Hype,Miggy Dela Rosa	2022-04-01	186661	90	deep groove house,edm,house,pop dance,tropical...	0.847	0.690	1	-7.877	0	0.0493	

Gambar IV- 9 Rekomendasi lagu

- Pemodelan Content-Based Filtering

Tahap pemodelan *content-based filtering* adalah pemodel sistem rekomendasi berdasarkan *content-based filtering*. Pada pemodelan ini kita dapat menentukan konten atau item yang ingin kita rekomendasikan berdasarkan *index* yang ingin digunakan. standar rekomendasi yang digunakan ialah *index* 0, kemudian akan menampilkan rekomendasi lagu berdasarkan *genre* yang mirip dengan *index* 0. Dapat dilihat sebagai berikut:

In [25]:	class RecommenderSystem: def __init__(self, data, content_col): self.df = pd.read_csv(data) self.content_col = content_col self.encoder = None self.bank = None self.df.isnull().sum() self.df.fillna(0).head() self.df = df.dropna(axis = 0)														
def fit(self): self.encoder = CountVectorizer(stop_words="english", tokenizer=word_tokenize) self.bank = self.encoder.fit_transform(self.df[self.content_col])															
def recommend(self, idx, top=5): content = df.loc[idx, self.content_col] code = self.encoder.transform([content]) dist = cosine_distances(code, self.bank) recommendation_ids = dist.argsort()[0, 0:(top+1)] return self.df.loc[recommendation_ids]															
In [26]:	recsys = RecommenderSystem("lagu_teratas_global.csv", content_col="Genres") recsys.fit()														
In [28]:	recsys.recommend(0)														
Out[28]:	Track Name	Album Name		Artist Name(s)	release	duration	Popularity	Genres	Danceability	Energy	Key	Loudness	Mode	Speechiness	Acous
0	Quevedo: Bzrp Music Sessions, Vol. 52	Quevedo: Bzrp Music Sessions, Vol. 52		Bizarrap,Quevedo	2022-07-06	198937	100	argentine hip hop,pop venezolano,trap argentin...	0.621	0.782	2	-5.548	1	0.0440	
20	STAYING ALIVE (feat. Drake & Lil Baby)	STAYING ALIVE (feat. Drake & Lil Baby)		Khaled,Drake,Lil Baby	2022-08-05	178176	85	dance pop,hip hop,miami hip hop,pop,pop,rap,ra...	0.718	0.461	4	-7.676	0	0.0797	
29	Bam (feat. Ed Sheeran)	Familia		Camila Cabello,Ed Sheeran	2022-04-07	206070	84	dance pop,pop,post-teen pop,uk pop	0.756	0.697	8	-6.377	1	0.0401	
19	About Damn Time	Special		Lizzo	2022-07-15	191937	80	dance pop,escape room,minnesota hip hop,pop,tr...	0.836	0.743	10	-6.305	0	0.0653	
14	Bad Decisions (with BTS & Snoop Dogg)	Bad Decisions (with BTS & Snoop Dogg)		benny blanco,BTS,Snoop Dogg	2022-08-05	172714	88	electropop,pop,rap,k-pop,k-pop boy group,g...	0.762	0.861	0	-3.296	1	0.0955	
35	Ferrari	Ferrari		James Hype,Miggy Dela Rosa	2022-04-01	186661	90	deep groove house,edm,house,pop dance,tropical...	0.847	0.690	1	-7.877	0	0.0493	

Gambar IV- 10 Pemodelan dan Pengujian data menggunakan Content-Based Filtering

4.2 Hasil

Berikut tabel lagu yang di rekomendasikan berdasarkan properti trek menggunakan *demographic filtering* dan *content-based filtering*.

Tabel IV- 2 Hasil

<i>Content-Based Filtering</i>
<i>Quevedo: Bzrp Music Sessions, Vol. 52</i>
<i>STAYING ALIVE (feat. Drake & Lil Baby)</i>
<i>Bam Bam (feat. Ed Sheeran)</i>
<i>About Damn Time</i>
<i>Bad Decisions (with BTS & Snoop Dogg)</i>
<i>Ferrari</i>

BAB V

PENUTUP

5.1.Kesimpulan

Berdasarkan hasil proses penelitian diatas, maka dapat diambil kesimpulan pada penelitian ini sebagai berikut:

- 1) 50 lagu tertas – global (*50 top songs - global*) dianalisa dengan cara melakukan *cleaning* data seperti: penyuntingan (*editing*), perumusan, pengujian data. Sistem rekomendasi digunakan dalam metode. *Content-Based Filtering* digunakan sebagai pemodelan dalam merekomendasikan lagu berdasarkan parameter dari properti trek serta menggunakan *Jupyter Notebook* untuk menampilkan hasil rekomendasi lagu berdasarkan properti trek.
- 2) Berdasarkan properti trek ‘genre’ terdapat 5 lagu teratas yang direkomendasikan menggunakan metode *Content-Based Filtering* yaitu *STAYING ALIVE (feat. Drake & Lil Baby)*, *Bam Bam (feat. Ed Sheeran)*, *About Damn Time*, *Bad Decisions (with BTS & Snoop Dogg)*, dan *Ferrari*.

5.2. Saran

Adapun saran untuk penelitian ini adalah hasil dari penelitian ini dapat menjadi bahan untuk penelitian selanjutnya yang lebih baik lagi adanya.

DAFTAR PUSTAKA

- Adhikari, S. (2019, February 28). *Building a Movie Recommendation Engine in Python using Scikit-Learn*. <https://medium.com/@sumanadadhikari/building-a-movie-recommendation-engine-using-scikit-learn-8dbb11c5aa4b>
- Fawzan, M. D., Setianingsih, C., & Nugrahaeni, R. A. (2021). Sistem Rekomendasi Lagu Metode Collaborative Filtering Berbasis Website Music Recomendation Sysystem Using Collaborative Filtering Filtering Website-Based. *E-Proceeding of Engineering*, 8, 12200–12206.
- Fitriani, S. A. (n.d.). *Analisis Klaster Atribut Musik pada Global Top 50 dengan Data Spotify dengan Menggunakan Algoritma K-Means*.
- Girsang, A. S. (2017, November 17). *Sistem rekomendasi- Content Based*. <https://mti.binus.ac.id/2020/11/17/sistem-rekomendasi-content-based/>
- K, R. C., & C, S. K. (2022). Switching Hybrid Model for Personalized Recommendations by Combining Users Demographic Information. *Journal of Theoretical and Applied Information Technology*, 15(3). www.jatit.org
- Kabiru, U., & Muhammad, A. (2021). Improving Personalized Questionnaire With Redundancy Reduction for Addressing Cold User Problem. *FUDMA JOURNAL OF SCIENCES*, 5(1), 457–466. <https://doi.org/10.33003/fjs-2021-0501-590>
- Kreitz, G., & Niemelä, F. (2010). Spotify – Large Scale, Low Latency, P2PMusic-on-Demand Streaming. *IEEE Communications Society Subject Matter Experts for Publication in the IEEE P2P 2010*.
- Liao, M., Sundar, S. S., & Walther, J. B. (2022, April 29). User Trust in Recommendation Systems: A comparison of Content-Based, Collaborative and Demographic Filtering. *Conference on Human Factors in Computing Systems - Proceedings*. <https://doi.org/10.1145/3491102.3501936>
- Liu, A., & Callvik, J. (2017). Using Demographic Information to Reduce the New User Problem in Recommender Systems. In *DEGREE PROJECT TECHNOLOGY*.
- Luo, K. (2018). *Machine Learning Approach for Genre Prediction on Spotify Top Ranking Songs*.
- Maringka, R., Khoirunnita, A., Maringka, R., & Utami, E. (2021). Analisa Perkembangan Musik Pada Spotify Menggunakan Structured Query Language (SQL) Analysis Of The Evolution Of Music On Spotify Using Structured Query Language (SQL). *Cogito Smart Journal |*, 7(1).
- Muslimin M. (n.d.). *Perkembangan Teknologi dalam Industri Media*.

- Narke, L., & Nasreen, A. (2020). A Comprehensive Review of Approaches and Challenges of a Recommendation System. *International Journal of Research in Engineering, Science and Management*, 3(4), 381–384. www.ijresm.com
- Noviani, D., Pratiwi, R., Silvianadewi, S., Alexandri, M. B., & Hakim, M. A. (2020). Pengaruh Streaming Musik Terhadap Industri Musik di Indonesia. In *Jurnal Bisnis STRATEGI* • (Vol. 29, Issue 1).
- Pereira, N., & Varma, S. L. (2019). Financial planning recommendation system using content-based collaborative and demographic filtering. *Advances in Intelligent Systems and Computing*, 669, 141–151. https://doi.org/10.1007/978-981-10-8968-8_12
- Roy, S., Sharma Professor, M., & Kumar Singh, S. (2019). Movie Recommendation System Using Semi-Supervised Learning. *Global Conference for Advancement in Technology (GCAT)*, 1–5.
- Sharma, D., Choudhury, T., Kumar Dewangan, B., Bhattacharya, A., & Dutta, S. (2021). *A Recommendation System for Customizable Items* (Vol. 1). <http://www.springer.com/series/11156>
- Softscients. (2022, April 5). *Cara Kerja Recommender Systems Continue reading at Cara Kerja Recommender Systems | Softscients*. https://softscients.com/2021/02/06/cara-kerja-recommender-systems/#Apa_itu_Recommender_Systems
- SpotifyAB. (2022). *Discover Spotify's FeaturesDiscover Spotify's Features*. <https://developer.spotify.com/documentation/web-api/reference/#/operations/get-audio-analysis>
- Stetler, R. (2022). Exploring Music Genres: A Study of Optimal Differentiation by Feature Feature. *Honors Projects*, 1–25. <https://scholarworks.bgsu.edu/honorsprojects>
- Toradmalle, D., Journal, I., Singh, T., Ladda, A., Shah, M., Davda, R., & Motekar, H. (2021). Recommendation System Using Hybrid Filtering. *International Research Journal of Engineering and Technology*. www.irjet.net
- Usino, W., Satria Prabuwono, A., Hamed Allehaibi, K. S., Bramantoro, A., & Amaldi, W. (2019a). Document Similarity Detection using K-Means and Cosine Distance. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 10, Issue 2). www.ijacsa.thesai.org
- Usino, W., Satria Prabuwono, A., Hamed Allehaibi, K. S., Bramantoro, A., & Amaldi, W. (2019b). Document Similarity Detection using K-Means and Cosine Distance. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 10, Issue 2). www.ijacsa.thesai.org
- Wahyu Dewatara, G., & Monik Agustin, S. (2019). Pemasaran Musik pada Era Digital Digitalisasi Industri Musik dalam Industri 4.0 di Indonesia. *WACANA: Jurnal Ilmiah Ilmu Komunikasi*, 18, 1–10.
- Wang, L.-D., Wang, X.-C., & Liu, K.-W. (2020). An Improved Multi-objective Shark Smell Optimization Algorithm for Automatic Train Operation Based on Angle Cosine and

Fusion Distance. *Journal of Computers*, 31(4), 141–156.
<https://doi.org/10.3966/199115992020083104011>

Zaidah, A. R., Septiarani, C. I., Nisa, S., Yusuf, A., & Wahyudi, N. (2021). *Komparasi Algoritma K-Means, K-Medoid, Agglomerative Clustering Terhadap Genre Spotify*. 7(1).
<https://ejurnal.fikom-unasman.ac.id>

Zuhraini, M., Gede, I., Suta Wijaya, P., & Bimantoro, F. (2022). *Temu Kembali Citra Menggunakan Metode Local Binary Pattern Rotation Invariant (Lbprot) dan Cosine Distance Similarity*. ARTICLE INFO ABSTRACT. 9(1), 1–10.
<https://dielektrika.unram.ac.id>

LAMPIRAN-LAMPIRAN

Lampiran 1. Bukti Kartu Bimbingan (KAMBING) dari Iteung



POLITEKNIK POS INDONESIA
PROGRAM STUDI D4 TEKNIK INFORMATIKA
JL. SARIASIH NO. 54 BANDUNG 40151
Telp. 022-2009562, 2009570
Fax. 022-2009568

FORMULIR KEGIATAN TUGAS AKHIR TA. 2021/2022

Nama : Naomi C.H Tampubolon
Npm : 1184018
Judul : Sistem Rekomendasi Lagu Teratas Global Berdasarkan Properti Trek Pengelompokan Menggunakan Metode Collaborative Filtering
Pembimbing : Rolly Maulana Awangga,, S.T., M.T.,CAIP.,SFPC



Pertemuan	Tanggal	Sudah Dikerjakan	Pekerjaan Selanjutnya	Nilai
1	04-08-2022	pembahasan topik pengembangan	mencari referensi topik yang akan diangkat	100
2	08-08-2022	bab 1	melanjutkan bab 2	100
3	09-08-2022	pembahasan bab 1	mengubah metodologi penelitian	100
4	10-08-2022	pembahasan metodologi	mencari referensi mengenai metode penelitian	100
5	11-08-2022	bab 2 dan 3	bab 4	100
6	12-08-2022	bab 4	implementasi yang akan diangkat	100
7	15-08-2022	laporan	jurnal	100
8	16-08-2022	demo aplikasi	demo aplikasi	70
			Rata-Rata:	96.25

Bandung, 06 September 2022

Pembimbing,



Rolly Maulana Awangga,, S.T., M.T.,CAIP.,SFPC
NIDN. 0410118609



POLITEKNIK POS INDONESIA
PROGRAM STUDI D4 TEKNIK INFORMATIKA
JL. SARIASIH NO. 54 BANDUNG 40151
Telp. 022-2009562, 2009570
Fax. 022-2009568

FORMULIR KEGIATAN

TUGAS AKHIR

TA. 2021/2022

Nama : Naomi C.H Tampubolon
Npm : 1184018
Judul : Sistem Rekomendasi Lagu Teratas Global Berdasarkan Properti Trek Pengelompokan Menggunakan Metode Collaborative Filtering
Pembimbing : Noviana Riza., S.Si., M.T.,SFPC



Pertemuan	Tanggal	Sudah Dikerjakan	Pekerjaan Selanjutnya	Nilai
1	01-08-2022	pembahasan topik	mencari referensi mengenai sistem rekomendasi dan topik yang berkaitan	90
2	02-08-2022	mencari referensi penelitian	menambahkan properti trek yang akan digunakan dan mencari referensi dan lanjut bab metodologi penelitian	90
3	03-08-2022	menambahkan properti trek yang akan digunakan metodologi penelitian dan mencari referensi	bab 123	86
4	06-08-2022	demo aplikasi	demo aplikasi	80
5	07-08-2022	demo aplikasi	demo aplikasi	80
6	08-08-2022	demo aplikasi	demo aplikasi	80
7	09-08-2022	demo aplikasi	demo aplikasi	80
8	16-08-2022	demo aplikasi	demo aplikasi	80
			Rata-Rata:	83.25

Bandung, 06 September 2022

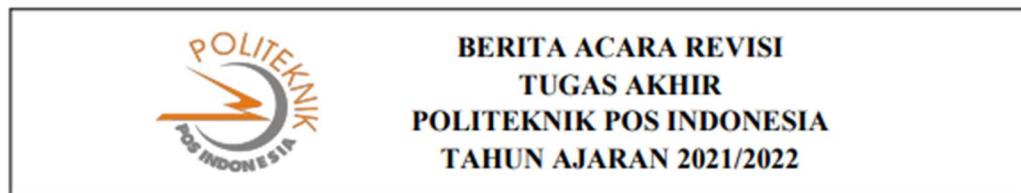
Pembimbing,



Noviana Riza., S.Si., M.T.,SFPC
NIDN. 0403117607

db7bb806-acef-467b-b50b-434265b02b6c

Lampiran 2. Bukti Berita Acara Revisi



Saya mahasiswa yang bertanda tangan sebagai berikut,

Nama Mahasiswa : Naomi C.H Tampubolon
NPM : 1184018
Judul : SISTEM REKOMENDASI LAGU TERATAS GLOBAL BERDASARKAN PROPERTI TREK PENGELOMPOKAN MENGGUNAKAN METODE COLLABORATIVE FILTERING
Tanggal : 30 Agustus 2022
Waktu : 14.00 - 15.00

telah melakukan revisi Tugas Akhir pada Prodi D4 Teknik Informatika Politeknik Pos Indonesia Tahun Ajaran 2021/2022 sesuai dengan peraturan yang berlaku, berdasarkan dengan perubahan aturan yang ditetapkan selama masa pandemi Covid 19 dengan kegiatan yang dilaksanakan secara daring dan dokumen elektronik.

Berikut hasil revisi Tugas Akhir:

Penguji Utama

No	Uraian	Tanda Tangan Mhs
1	Buku	

Penguji Pendamping

No	Uraian	Tanda Tangan Mhs
1	Menambahkan popularity dalam proses data menggunakan model demographic filtering	
2	Melengkapi penjelasan dan langkah-langkah model demographic	

Bandung, 30 Agustus 2022

Mengetahui:

Penguji Utama

Penguji Pendamping



(Rolly Maulana Awangga, S.T., M.T.,CAIP.,SFPC) (Nisa Hanum Harani, S.Kom., M.T.,CDSP.,SFPC)

Diperiksa oleh,

Ketua
Program Studi D4 Teknik Informatika



Koordinator Tugas Akhir
Program Studi D4 Teknik Informatika



(M. Yusril Helmi S, S.Kom., M.Kom.)

(Roni Andarsyah, ST., M.KOM.,SFPC)

dc365499-0a28-4b68-8f09-f187a94e49e8

Lampiran 3. Bukti Pengecekan Plagiarisme Online

The screenshot shows a plagiarism checker interface with a table of results and a search result for BSV Academy.

PLAGIARISM CHECKER

0% Unique Content

#	String	Uniqueness
1	beberapa unsur, seperti melodi, harmoni, ritme, dan timbre.	Already Exists
2	harmoni, ritme, dan timbre. Musik adalah fenomena intuitif	Already Exists
3	dikembangkan, ditingkatkan, dan disebarluaskan dengan	Already Exists
4	berbagai cara. Musik adalah fenomena unik yang dihasilkan	Already Exists
5	berbagai alat musik. Musik memang menjadi salah satu hobi yang	Already Exists
6	paling digemari, hal tersebut terbukti dengan manfaat yang	Already Exists
7	mendengarkan musik. Musik memiliki kemampuan untuk mengurangi	Already Exists
8	mengurangi stres. Manfaat mendengarkan musik adalah kita bisa	Already Exists
9	melakukan hal lain. Musik adalah komponen yang tak terpisahkan	Already Exists
10	dan teknologi. Perkembangan yang semakin berkembang pesat	Already Exists
11	dunia industri.(Muslimin M, n.d.) Dimana pada jaman era	Already Exists
12	dan kapanpun, dikarenakan kecanggihan teknologi tersebut.	Already Exists
13	teknologi tersebut. Spotify adalah layanan streaming musik terpopuler	Already Exists
14	di Indonesia, telah mencapai 200 juta pengguna perbulan.(Noviani	Already Exists
15	pengguna perbulan.(Noviani et al., 2020) Data yang digunakan	Already Exists
16	Songs - Global).(Fitriani, n.d.) Spotify memiliki properti	Already Exists
17	sebuah musik. Ada beberapa fitur atau properti lagu pada	Already Exists
18	spotify yaitu, sebagai berikut: track name, album name,	Already Exists

https://www.googleadservices.com/pagead/ack?sa=L&ai=Czpp5X8cXY_NeD4OVstlPypSJu...

SEARCH

Search from seo tools

BSV ACADEMY

EARN YOUR
Introduction to
BSV Development
CERTIFICATE

ENROL NOW FOR FREE

Popular SEO Tools

Lampiran 4. Surat Pernyataan

SURAT PERNYATAAN TIDAK MELAKUKAN PLAGIARISME

Yang bertanda tangan di bawah ini:

Nama : Naomi C.H Tampubolon
NPM : 1184018
Program Studi : D4 Teknik Informatika
Judul : Sistem Rekomendasi Lagu Teratas - Global Berdasarkan Properti Trek Menggunakan Metode *Content-Based Filtering*

Menyatakan bahwa:

1. Program Tugas Akhir Saya ini adalah asli dan belum pernah diajukan untuk memenuhi kelulusan matakuliah Tugas Akhir pada Program Studi D4 Teknik Informatika baik di Politeknik Pos Indonesia maupun di Perguruan Tinggi lainnya.
2. Program Tugas Akhir ini adalah murni gagasan, rumusan, dan penelitian saya sendiri tanpa bantuan pihak lain, kecuali arahan pembimbing.
3. Dalam Program Tugas Akhir ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dicantuman sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan-penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya saya ini, serta sanksi lainnya sesuai dengan norma yang berlaku diperguruan tinggi lain.

Bandung, 30 Agustus 2022
Yang membuat pernyataan,

Naomi C.H Tampubolon
NPM. 1.18.4.018

Lampiran 5. Draft Jurnal

Rolly et al. / Jurnal Sistem Cerdas (20XX) Vol XX- No XX

Sistem Rekomendasi Lagu Teratas – Global Berdasarkan Properti Trek menggunakan Metode Content Based Filtering

Rolly Maulana Awangga
Jurusan Teknik Informatika
Politeknik Pos Indonesia
Bandung, Indonesia
awangga@poltekpos.ac.id

Naomi C.H Tampubolon*
Jurusan Teknik Informatika
Politeknik Pos Indonesia
Bandung, Indonesia
naomicindi.hermina@gmail.com

Spotify saat ini merupakan platform musik, podcast dan video yang banyak digemari banyak penggunanya. Spotify juga memberikan kemudahan akses musik. Platform musik spotify sangat banyak digemari penggunanya, terutama kalangan young adults yaitu berusia 18 hingga 25 tahun.(Wahyu Dewatara & Monik Agustin, 2019) Spotify menyediakan banyak fitur seperti yaitu sekarang memutar, cari, koleksi, dengarkan offline, podcast dan acara serta banyak lagi. Properti trek di Spotify termasuk genre, nama artis, dan banyak lainnya. (Fitriani, n.d.) Berdasarkan properti trek, bagaimana cara merekomendasikan lagu berdasarkan properti trek, dan bagaimana cara memastikan bahwa lagu berdasarkan properti trek disukai. Masalah ini akan menjadi sumber dalam penelitian. *Dataset* yang digunakan dalam penelitian ini adalah 50 Lagu Teratas – Global (*50 Top Songs - Global*) yang diambil dari Spotify playlist chart yang kemudian diconvert ke dalam dalam file dengan format csv dan kemudian nantinya akan dikelola atau diproses. Penelitian ini bertujuan untuk merekomendasikan lagu berdasarkan trek properti yang ada dan merekomendasikan berdasarkan trek properti disukai. Metode Content-Based Filtering digunakan dalam penelitian ini. *Content-Based Filtering* berfungsi merekomendasikan berdasarkan kemiripan konten seperti genres yang merupakan kemiripan jarak cosinusnya dan terdapat 5 lagu teratas yaitu *STAYING ALIVE (feat. Drake & Lil Baby)*, *Bam Bam (feat. Ed Sheeran)*, *About Damn Time, Bad Decisions (with BTS & Snoop Dogg)*, dan *Ferrari* yang direkomendasikan menggunakan metode *Content-Based Filtering*.

Kata Kunci: *Sistem rekomendasi, musik, spotify, content-based filtering*.

I. PENDAHULUAN

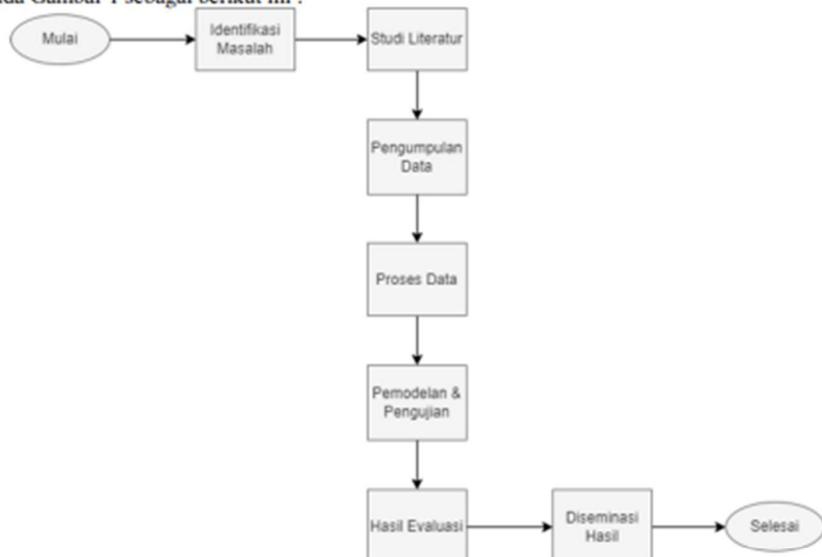
Musik terdiri dari beberapa unsur, seperti melodi, harmoni, ritme, dan timbre. Musik adalah fenomena *intuitif* yang dapat dikembangkan, ditingkatkan, dan disebarluaskan dengan berbagai cara. Musik adalah fenomena unik yang dihasilkan oleh berbagai alat musik. Musik memang menjadi salah satu hobi yang paling digemari, hal tersebut terbukti dengan manfaat yang bisa didapatkan dengan mendengarkan musik. Musik memiliki kemampuan untuk mengurangi stres. Manfaat mendengarkan musik adalah kita bisa mendengarkannya sambil melakukan hal lain. Musik adalah komponen yang tak terpisahkan dari kehidupan dan teknologi. Perkembangan yang semakin berkembang pesat juga mempengaruhi dunia industri.(Muslimin M, n.d.) Dimana pada jaman era digital saat ini tidak perlu repot dalam menggengarkan musik dimanapun dan kapanpun, dikarenakan kecanggihan teknologi tersebut. Spotify adalah layanan *streaming* musik terpopuler di Indonesia, telah mencapai 200 juta pengguna perbulan.(Noviani et al., 2020) Data yang digunakan pada penelitian ini adalah data lagu yang diambil dari *playlist chart spotify* 50 Lagu Teratas – Global (*50 Top Songs - Global*). (Fitriani, n.d.) Spotify memiliki properti lagu yang berfungsi untuk mengetahui lebih mendalam mengenai lagu atau sebuah musik. Ada beberapa fitur atau properti lagu pada spotify yaitu, sebagai berikut: *track name, album name, artist name, duration, popularity, genres, danceability, energi, key, loudness, speechiness, acousticness, instrumentaness, liveness, tempo* dan *valence*. Berdasarkan properti trek tersebut kita dibutuhkan sistem yang dapat merekomendasikan berdasarkan properti trek yang ada. Maka dari itu bagaimana caranya untuk merekomendasikan sebuah lagu berdasarkan properti trek yang ada. Dari permasalahan tersebut dibutuhkan sebuah

sistem yang dapat merekomendasikan lagu berdasarkan properti trek yang ada. Properti trek 'genres' merupakan properti trek yang digunakan dalam rekomendasi sistem.

Dalam penelitian ini sistem rekomendasi akan dianalisa dan dibangun dengan menggunakan metode *content-based filtering* karena dengan algoritma ini berdasarkan *similarity* konten sehingga rekomendasi yang dihasilkan mempunyai kualitas yang baik. Visualisasi hasil rekomendasi tersebut akan diimplementasikan dengan *framework Jupyter Notebook*. Target luaran yang ingin dicapai pada penelitian ini adalah jurnal nasional terakreditas.

II. METODOLOGI PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini akan diwakilkan oleh flowchart pada Gambar 1 sebagai berikut ini :



Gambar 1. Diagram Alur Metodologi Penelitian

A. Identifikasi Masalah

Identifikasi masalah merupakan langkah pertama dalam tahapan metode penelitian yang berfungsi untuk mengidentifikasi masalah. Pada tahap ini peneliti melakukan penelitian dengan cara memunculkan masalah (*problem*) yang ada, kemudian melakukan indentifikasi sumber permasalahan (*root cause*) sehingga menciptakan permasalahan (*problem statement*) yang menjelaskan permasalahan yang sudah diidentifikasi. Adapun Permasalahan yang terdapat pada penelitian ini adalah bagaimana cara merekomendasikan lagu berdasarkan properti trek.

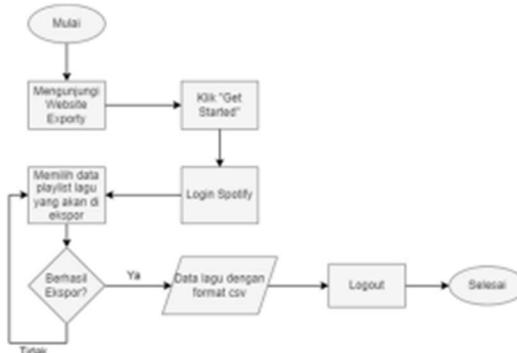
B. Studi Literatur

Setelah dilakukan tahap indentifikasi masalah langkah selanjutnya dalam tahapan metode penelitian adalah studi literatur. Pada tahap ini peneliti melakukan pencarian informasi yang relevan dengan masalah yang akan diteliti dan mengkaji beberapa teori dasar yang relevan dengan permasalahan yang akan diteliti kemudian memperdalam pengetahuan peneliti tentang permasalahan yang akan diteliti dan mengkaji hasil-hasil penelitian terdahulu yang berkaitan dengan penelitian yang akan diteliti dan mendapatkan

informasi tentang aspek-aspek mana dari suatu permasalahan yang sudah pernah diteliti untuk menghindari agar tidak terjadi *plagiarisme* penelitian.

C. Pengumpulan Data

Pengumpulan data adalah tahap setelah studi literatur. Tujuan pengumpulan data adalah untuk memperoleh data yang akan digunakan dalam penelitian. Peneliti melakukan pengumpulan studi literatur yang relevansi dan sesuai dengan apa yang dibutuhkan untuk menunjang penelitian untuk dapat melakukan teknik pengumpulan data studi literatur. Maka dari itu data yang digunakan pada penelitian ini adalah data yang di *convert* dari website <https://exportify.net/#playlists> kedalam bentuk *format csv*. Website tersebut berfungsi untuk mengekspor daftar putar *spotify* menggunakan *Web API*. Alur dari pengumpulan data dapat dilihat dari *flowchart* berikut:



Penjelasan *flowchart*, dimulai dengan cara mengunjungi website Exportify kemudian akan muncul tampilan awal dari website. Klik “Get Started”, kemudian akan dialihkan ke halaman login dan melakukan proses login. Setelah masuk login selanjutnya akan di tampilkan daftar spotify yang sudah ditambahkan kedalam koleksi playlist pada spotify. Kemudian klik export lalu playlist yang berisi lagu yang ingin diambil datanya akan terekspor menjadi data yang berbentuk file csv. kemudian logout setelah selesai melakukan export data.

D. Proses Data

Setelah dilakukan pengumpulan data langkah selanjutnya dalam tahap metodologi penelitian adalah proses data. Pada tahap proses data atau pengolahan data dilakukan data *cleaning*. *Cleaning* data adalah prosedur untuk memastikan keakuratan, konsistensi, dan kegunaan data dalam kumpulan data. Prosedurnya adalah mendeteksi kesalahan atau kerusakan pada data dan kemudian memperbaiki atau menghapusnya jika perlu.

E. Pemodelan dan Pengujian

Tahapan selanjutnya yaitu Pemodelan dan Pengujian. Tahap ini ialah tahap menggunakan metode rekomendasi dalam pemodelan dan pengujian data. *Content-Based Filtering* merupakan metode, model atau algoritma yang akan digunakan pada penelitian ini. Pemodelan ada tahap yang dilakukan sebelum pengujian. Dalam pemodelan dipertimbangkan dulu model yang akan dipakai dalam mengelola data. Kemudian langkah selanjutnya setelah pemodelan adalah langkah Pengujian menggunakan model *Content-Based Filtering*.

F. Hasil Evaluasi

Tahap hasil evaluasi adalah tahap performansi dari pemodelan dan pengujian. Hasil pengujian data akan diimplementasi pada tahap hasil evaluasi.

G. Diseminasi Hasil

Tahap deseminasi hasil adalah tahap untuk artikel yang diterbitkan dalam laporan penelitian dan *draft* jurnal.

III. PEMBAHASAN DAN HASIL

A. PEMBAHASAN

Penelitian ini dilakukan dengan mengambil data selama 1 minggu terakhir dari *playlist* 50 Lagu Teratas - Global (*50 Top Songs - Global*) sebanyak 50 data lagu. Untuk hasil data mentah atau data yang belum dilakukan proses data yaitu data yang ada pada *spotify*, dapat dilihat sebagai berikut:

Rank	Artist	Title	Plays	Album	Length
1	Quevedo Bago Music Sessions, Vol. 82	Blanca, Quevedo	49,993,772	Quevedo Bago Music Sessions, Vol. ...	3:19
2	As It Was	Harry Styles	28,156,789	Harry's House	2:47
3	Me Perdi Bonita	Elvi Bunny, Chencito Corleone	24,028,045	Un Verano Sin Ti	2:58
4	Tú Me Preguntás	Elvi Bunny	20,151,875	Un Verano Sin Ti	4:05
5	Raining Up That Hill (A Deal With God)	Kate Bush	20,032,750	Hounds Of Love	4:58
6	DESPICHA	ROBALIA	20,307,848	DESPICHA	2:27
7	Él	Bad Bunny	20,359,490	Un Verano Sin Ti	2:33
8	Glimpse of Us	Jay	20,333,979	Glimpse of Us	2:55
9	Ojitos Lindos	Bad Bunny, Banda Estéreo	23,498,678	Un Verano Sin Ti	4:38
10	Bad Habit	Steve Lacy	21,208,876	Genius Rights	2:02
11	Moscow Mule	Bad Bunny	20,846,139	Un Verano Sin Ti	4:06
12	Heat Waves	Glass Animals	20,721,676	Dreamland (+ Bonus Levels)	2:00
13	La Rosalita	Monica Santa	20,183,692	La Rosalita	2:42
14	PROVENZA	KAROL G	19,745,406	PROVENZA	3:30
15	Bad Decisions (with BTS & Snoop Dogg)	BTS, Snoop Dogg	19,500,644	Bad Decisions (with BTS & Snoop ...)	2:52
16	I Ain't Worried	OneRepublic	18,361,468	I Ain't Worried (Music From The M...)	2:09

17	 Left and Right (Feat. Jung Kook of BTS) Charlie Puhn, Jung Kook, BTS	18,278,990	Left and Right (Feat. Jung Kook of ...	0:34
18	 Deja Central Cee	17,242,853	Deja	1:27
19	 BREAK MY SOUL Beyond	16,062,071	RENAISSANCE	4:38
20	 About Damn Time Lil Nas X	15,981,437	Spectre	5:11
21	 STAYING ALIVE (Feat. Drake & Lil Baby) Drake, Lil Baby	15,993,590	STAYING ALIVE (Feat. Drake & Li ...	2:58
22	 STAY (With Justin Bieber) The Kid LAROI, Justin Bieber	15,647,793	F*CK LOVE IS OVER YOU	0:21
23	 Taxid Bad Bunny, Jhay Cortez	15,756,452	Un Verano Sin Ti	5:57
24	 Cold Heart - PNAU Remix Elon John, Dua Lipa, PNAU	15,285,197	The Lockdown Sessions	5:02
25	 Te Felicito Shakira, Juan Alejandro	15,230,719	Te Felicito	0:52
26	 RUM Bella Pozen	15,018,493	SNAP PACK	0:59
27	 Late Night Talking Harry Styles	14,395,673	Harry's House	0:57
28	 Party Bad Bunny, Juan Alejandro	14,546,290	Un Verano Sin Ti	5:47
29	 I Like You (A Happier Song) (With Doja Cat) Post Malone, Doja Cat	14,533,671	Twelve Carat Toothache	3:12
30	 Bam Bam (Feat. Ed Sheeran) Camila Cabello, Ed Sheeran	14,038,418	Pendula	3:06
31	 Jenny from the Block (Feat. 21 Savage) Drake, 21 Savage	13,188,679	Honesty, Nevermind	3:38
32	 Under the Influence Chris Brown	13,143,793	Indigo (Deluxe)	3:04
33	 Sunroof Nicky Jam, Daddy Yankee	13,141,181	Sunroof	2:43
34	 Sweater Weather The Neighbourhood	13,607,763	I Love You	4:00
35	 Blinding Lights The Weeknd	12,478,897	After Hours	3:30
36	 Ferrari James Hope, Migos, Doja Cat	12,326,457	Perreo	3:06
37	 Ghost Justin Bieber	12,021,064	Justice	3:33
38	 La Cortejante Bad Bunny, Tini	12,329,312	Un Verano Sin Ti	3:38
39	 One Kiss (With Dua Lipa) Calvin Harris, Dua Lipa	11,851,800	One Kiss (With Dua Lipa)	3:34
40	 Vegas (From the Original Motion Picture Soundtrack ...) Doja Cat	11,932,257	Vegas (From the Original Motion P...)	3:02
41	 Ole For You The Weeknd	11,881,548	Barney	4:23
42	 Until I Found You Stephen Sanchez	11,871,498	Until I Found You	2:57
43	 MIDDLE OF THE NIGHT Elley Duhé	11,830,079	MIDDLE OF THE NIGHT	3:04
44	 First Class Jack Harlow	11,457,292	Come Home The Kids Miss You	3:53
45	 Another Love Tom Odell	11,480,354	Leng Way Down (Deluxe)	4:04
46	 Una Noche en Medellin Cris M	11,180,078	Una Noche en Medellin	3:53
47	 Kesiya (From 'Brahmastra') Pritam, Arijit Singh, Amitabh Bhattacharya	10,904,301	Kesiya (From 'Brahmastra')	4:08
48	 Despide de la Playa Bad Bunny	10,825,650	Un Verano Sin Ti	5:50
49	 Candela Rauw Alejandro	10,658,440	Safe Haven	5:53
50	 Shivers Ed Sheeran	10,503,440	-	3:07

Kemudian data diconvert dalam format file csv.

Spotify ID	Artist IDs	Track Name	Album Name	Artist Name(s)	Release Date
2tTmW7RDTMQtBk7m2rYeSw	"716NhGYqD1j2wl1Qkgq36,52iwsT98xCoGgiG				
4Dvkj6hhA12EX05fT7y2e	"6KlmCVd70vtloJWnq6nGn3"	"As It Was"	"Harry's		
6sq7ltF9Qa7SNFBsV5Cogx	"4q3ewBCX7sLwd24euuV69X,37230BxxYs9ksS7O1				
1IHWi5LamUGEuP4o2kQXSZ	"4q3ewBCX7sLwd24euuV69X"	"Tití Me Preguntó"			
75FEaRjZTKLhTrFGsfMUXR	"1aSxMhuvisZ8h9dk9jlDwL"	"Running Up That Hill"			
5ildQOEKmJuWGl2vRkfDyc	"7ltDVBr6mKbRvohxheJ9h1"	"DESPECHÁ"	"DESPE		
5Eax0qFko2dh7RI2lYs3bx	"4q3ewBCX7sLwd24euuV69X"	"Efecto"	"Un Verano		
6xGruZOHLs39ZbVccQTuPZ	"3MZsBdqDrTJihTHQrO6Dq"	"Glimpse of Us"	"G		
3k3NWokhRRkEPheCzPrmV8TW	"4q3ewBCX7sLwd24euuV69X,5n9bMYfz9qss2"				

Proses selanjutnya dalam penelitian ini adalah proses pemodelan dan pengujian data. Penelitian ini menggunakan *content-based filtering* dalam pemodelannya. Langkah-langkah dalam pemodelan menggunakan *content-based filtering*.

a) *Import Library*

```
In [1]: import pandas as pd  
import numpy as np  
import sklearn  
import nltk  
from sklearn.metrics.pairwise import cosine_distances  
from sklearn.feature_extraction.text import CountVectorizer
```

b) Menampilkan 5 dataset teratas

1. Tampilkan 3 dataset teratas.

Global Music Sales Performance - Q3 2023								
Song ID	Song Details		Artist & Track Info		Sales & Distribution			
	Title	Artist	Name	Type	Label	Format	Region	Popularity
S-001	Eye of the Tiger	Rick Astley	Eye of the Tiger	Single	Universal Music Group	Digital Download	North America	Very High
S-002	Hotel California	The Eagles	Hotel California	Single	Columbia Records	Digital Download	Europe	High
S-003	Billie Jean	Michael Jackson	Billie Jean	Single	Sony Music	Digital Download	Asia Pacific	Very High
S-004	Eye of the Tiger	Rick Astley	Eye of the Tiger	Single	Universal Music Group	Digital Download	North America	Very High
S-005	Hotel California	The Eagles	Hotel California	Single	Columbia Records	Digital Download	Europe	High
S-006	Billie Jean	Michael Jackson	Billie Jean	Single	Sony Music	Digital Download	Asia Pacific	Very High
S-007	Eye of the Tiger	Rick Astley	Eye of the Tiger	Single	Universal Music Group	Digital Download	North America	Very High
S-008	Hotel California	The Eagles	Hotel California	Single	Columbia Records	Digital Download	Europe	High
S-009	Billie Jean	Michael Jackson	Billie Jean	Single	Sony Music	Digital Download	Asia Pacific	Very High
S-010	Eye of the Tiger	Rick Astley	Eye of the Tiger	Single	Universal Music Group	Digital Download	North America	Very High
S-011	Hotel California	The Eagles	Hotel California	Single	Columbia Records	Digital Download	Europe	High
S-012	Billie Jean	Michael Jackson	Billie Jean	Single	Sony Music	Digital Download	Asia Pacific	Very High

8 rows x 23 columns

Cleaning Data
Cleaning data adalah prosedur untuk memastikan keakuratan, konsistensi, dan kegunaan data dalam kumpulan data. Prosedurnya adalah mendeteksi kesalahan atau kerusakan pada data dan kemudian memperbaiki atau menghapusnya jika perlu. Berikut adalah proses data yang telah dilakukan *cleaning* data:

periode. Berikut adalah:

Drop tabel
Drop tabel yang befungsi untuk menghilangkan tabel yang tidak digunakan atau diperlukan dalam proses penelitian. Terdapat beberapa data yang tidak diperlukan dalam penelitian ini, dan di lakukan *drop tabel*. Dapat dilihat sebagai berikut:

In [4]: `df[0].head(10).T.to_csv('test.txt', index=False)`

3-2021-09

Rank	Track Name	Album Name	Artist Name(s)	Release Date	Duration	Popularity	Score	Censorship	Energy	Key	Loudness	Mode	Sp
1	Granite Bedroom Songbook Vol. Vol. 12	Granite Bed- room Songbook Vol. Vol. 12	Broadway Quartet	2020-07-26	168:07	100	argentine	arg	0.824	0.762	2	-5.548	1
2	It's All Yours	Harry's House	Harry Styles	2020-07-26	167:03	93	pop	0.826	0.751	6	-5.338	0	
3	We Don't Talk Anymore	Un Verano Sin U.S.	Bad Bunny/ Cazzu	2018-06-28	170:07	95	www.vevo.com/watch/bad-bunny-we-dont-talk-anymore	0.841	0.772	1	-5.105	0	
4	Tell Me About It	Un Verano Sin U.S.	Bad Bunny	2020-06-28	242:16	90	www.vevo.com/watch/bad-bunny-tell-me-about-it	0.838	0.716	5	-5.188	0	
5	Running In Her Circle	Hours Of Hell	Kate Bush	1980-01-01	260:03	90	art pop art rock business pop rock business pop rock post punk	0.826	0.847	10	-5.125	0	

- Cek Data

Mengecek apakah data tersebut memiliki *missing value* atau tidak. Setelah dilakukan pengecekan data terdapat 2 data yang memiliki *missing value*.*Drop* tabel yang tidak digunakan.

Lalu dilakukan proses mengisi nilai *NA/Nan (Not a Number)* menjadi 0.

Index	Song ID	Title	Artist Name(s)	Release Date	Duration	Popularity	Genres	Danceability	Energy	Key	Loudness	Mode	Speechiness	Acousticness
0	SG-001	Guacamole	Queens of the Stone Age	2002-07-08	196007	100	alternative hip hop	0.621	0.762	2	-5.546	1	0.040	0.8
1	AS-002	As It Was	Harry Styles	2020-08-26	167303	30	pop	0.820	0.731	8	-6.326	0	0.007	0.3
2	MP-003	My Darlin' Bombs	U2	1980-05-25	175007	30	alternative rock	0.411	0.712	1	-4.165	0	0.007	0.6
3	TB-004	To Be Your Lover	Blink-182	2000-09-08	242718	30	alternative rock	0.400	0.710	9	-6.190	0	0.030	0.6
4	BU-005	Running Up That Hill	Hounds Of Love	1986-10-06	206612	30	art pop art rock	0.820	0.647	10	-13.120	0	0.090	0.7

Kemudian menghapus atau menghilangkan semua entri dengan nilai *NaN* atau 0 seperti berikut:

In [30]: `df = df.dropna(inplace=True)`

- d) Pemodelan sistem rekomendasi berdasarkan konten

Pada pemodelan menggunakan *content-based filtering* memiliki beberapa tahap dalam menampilkan rekomendasi berdasarkan konten. Dapat dilihat sebagai berikut:

- *Encode Data*

Pada tahap ini `genres` di `encode` dengan cara mengkonversi `teks` menjadi `vector` menggunakan `CountVectorizer` dengan konten yang digunakan dalam merekomendasikan yaitu kemiripan konten index 0 pada properti `trek genres`. Dengan langkah sebagai berikut:

content-based filtering

- Mencari *similarity*

Pada tahap ini kita menggunakan *cosine distance* dalam mencari kesamaan atau kemiripan konten dengan cara menghitung jarak *cosinus* yaitu ukuran

kesamaan antara dua vektor berdasarkan sudut kosinus di antara konten. (Usino et al., 2019) Kemudian mengembalikan nilai index bilangan bulat yang akan mengurutkan indeks. Dari hasil mengurutkan tersebut maka didapat hasil rekomendasi berdasarkan *index* 0 yaitu 5 lagu teratas yang mirip dengan konten yang sama dengan *index* 0 dan dapat dilihat sebagai berikut:

- Pemodelan *Content-Based Filtering*.

Tahap pemodelan *content-based filtering* adalah pemodel sistem rekomendasi berdasarkan *content-based filtering*. Pada pemodelan ini kita dapat menentukan konten atau item yang ingin kita rekomendasikan berdasarkan *index* yang ingin digunakan. standar rekomendasi yang digunakan ialah *index* 0, kemudian akan menampilkan rekomendasi lagu berdasarkan *genre* yang mirip dengan *index* 0. Dapat dilihat sebagai berikut:

```

In [18]: class RecommenderSystem:
    def __init__(self, data, content_col):
        self.data = pd.read_csv(data)
        self.content_col = content_col
        self.title_col = "title"
        self.rank = 5000
        self.df_content = self.data[[self.content_col, "score"]]
        self.df_titles = self.data[[self.title_col, "count"]]
        self.df_titles["tf_idf"] = tf_idf(self.title_col)
        self.df_titles["tf_idf_rank"] = self.rank * self.df_titles["tf_idf"]

    def fit(self):
        self.encoder = CounterVectorizer(stop_words="english", token_ngram_range=(1, 2))
        self.bow = self.encoder.fit_transform(self.data[self.content_col])

    def recommend(self, title, top_k=5):
        content_ids = np.where(self.content_col == title)[0]
        content_ids = np.append(content_ids, np.where(self.title_col == title)[0])
        dists = cosine_distances(content_ids, self.bow)
        recommendation_ids = dist.argsort(0, top_k)
        return self.df_titles[recommendation_ids]

In [19]: recommender = RecommenderSystem("labeled_data.csv", content_col="titles")
recommender.fit()

```

DRAFT [28]: PREDICT RECOMMENDATION														
Rank	Track Name	Album Name	Artist Name(s)	Release Date	Duration	Popularity	Danceability	Conducive	Energy	Key	Loudness	Mode	Speechiness	Acoustic
9	Genetics	Quevedo: Bzrp Music Sessions, Vol. 52	Quevedo	2020-07-08	100:00:07	100	argentine hip-hop reggaeton	0.621	0.762	2	-5.540	+	0.3440	
20	STAYING ALIVE (feat. Drake & Lil Baby)	STAYING ALIVE	Khalid, Drake, LF Baby	2020-09-05	17:57:15	100	dance-pop trap trap-reggaeton	0.718	0.411	4	-7.676	0	0.3797	
28	Bam Bam (feat. Ed Sheeran)	Panda	Gentle Cancer (ft. Sheeran)	2020-04-07	20:00:05	100	dance-pop trap dance-trap	0.768	0.697	6	-6.377	1	0.3401	
19	About Damn Time	Spaceman	U2	2020-07-10	10:58:07	100	dance-pop trap rock-trap	0.638	0.143	10	-6.395	0	0.3950	
14	Bad Decisions (with BTS & Snoop Dogg)	Bad Decisions (with BTS & Snoop Dogg)	Jenny Lewis, BTS, Snoop Dogg	2020-09-03	17:27:14	100	alternatives-punk rock rap & trap-trap-folk	0.710	0.861	0	-3.298	1	0.3568	
36	Ferrari	Ferrari	James Mingo Jones	2020-04-01	10:00:01	100	house-electro-trap dance-trap	0.647	0.690	4	-7.677	0	0.3400	

B. HASIL

Berikut tabel lagu yang di rekomendasikan berdasarkan properti trek menggunakan *content-based filtering*.

Content-Based Filtering
<i>Quevedo: Bzrp Music Sessions, Vol. 52</i>
<i>STAYING ALIVE (feat. Drake & Lil Baby)</i>
<i>Bam Bam (feat. Ed Sheeran)</i>
<i>About Damn Time</i>
<i>Bad Decisions (with BTS & Snoop Dogg)</i>
<i>Ferrari</i>

IV. KESIMPULAN DAN SARAN

A. KESIMPULAN

Berdasarkan hasil proses penelitian diatas, maka dapat diambil kesimpulan pada penelitian ini sebagai berikut:

- 50 lagu tertas – global (*50 top songs - global*) dianalisa dengan cara melakukan *cleaning* data seperti: penyuntingan (*editing*), perumusan, pengujian data. Sistem rekomendasi digunakan dalam metode. *Content-Based Filtering* digunakan sebagai pemodelan dalam merekomendasikan lagu berdasarkan parameter dari properti trek serta menggunakan *Jupyter Notebook* untuk menampilkan hasil rekomendasi lagu berdasarkan properti trek.
- Berdasarkan properti trek ‘genre’ terdapat 5 lagu teratas yang direkomendasikan menggunakan metode *Content-Based Filtering* yaitu *STAYING ALIVE (feat. Drake & Lil Baby)*, *Bam Bam (feat. Ed Sheeran)*, *About Damn Time*, *Bad Decisions (with BTS & Snoop Dogg)*, dan *Ferrari*.

B. SARAN

Adapun saran untuk penelitian ini adalah hasil dari penelitian ini dapat menjadi bahan untuk penelitian selanjutnya yang lebih baik lagi adanya.

REFERENCES

- Fitriani, S. A. (n.d.). *Analisis Klaster Atribut Musik pada Global Top 50 dengan Data Spotify dengan Menggunakan Algoritma K-Means*.
Muslimin M. (n.d.). *Perkembangan Teknologi dalam Industri Media*.

- Noviani, D., Pratiwi, R., Silvianadewi, S., Alexandri, M. B., & Hakim, M. A. (2020). Pengaruh Streaming Musik Terhadap Industri Musik di Indonesia. In *Jurnal Bisnis STRATEGI* • (Vol. 29, Issue 1).
- Usino, W., Satria Prabuwono, A., Hamed Allehaibi, K. S., Bramantoro, A., & Amaldi, W. (2019). Document Similarity Detection using K-Means and Cosine Distance. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 10, Issue 2). www.ijacsa.thesai.org
- Wahyu Dewatara, G., & Monik Agustin, S. (2019). Pemasaran Musik pada Era Digital Digitalisasi Industri Musik dalam Industri 4.0 di Indonesia. *WACANA: Jurnal Ilmiah Ilmu Komunikasi*, 18, 1–10.