

## Analisa Eksplisit Pada Spotify Berbasis Big Data

### *Explicit Analysis of Spotify Based on Big Data*

**Yoga Putra Pratama<sup>1\*</sup>, Heribertus Rudi Kusumantoro<sup>2</sup>**

<sup>1,2</sup>Politeknik Negeri Jakarta

Jl. Prof. DR. G.A. Siwabessy, Kampus Universitas Indonesia Depok, 16425, 021-7270036

yoga.putra.pratama@grafika.pnj.ac.id<sup>1\*</sup>, rudi.kusumantoro@grafika.pnj.ac.id<sup>2</sup>

***Abstraks-** Internet merupakan bagian yang tidak terpisahkan dalam kehidupan saat ini. Banyak aplikasi yang menggunakan internet untuk diakses dalam kehidupan seperti salah satunya yaitu SPOTIFY, Spotify merupakan platform musik online yang terdapat banyak daftar lagu maupun musisi didunia. Namun pada spotify terdapat jenis lagu eksplisit yaitu lagu yang didalam liriknya terdapat kata kata umpatan, kasar atau diskriminatif. Big data digunakan untuk analisis data dalam aplikasi Snowflake. Untunya, k mengetahui dibuat daftar terbaru music dan Musisi yang menggunakan kalimat eksplisit dengan mengintegrasikan data yang bersumber dari Kaggle berupa CSV. Setelah dilakukan integrasi maka didapatkan hasil lagu ataupun Musisi yang menggunakan kata eksplisit bertujuan agar mengetahui isi konten dari lagu tersebut.*

***Kata kunci:** eksplisit,spotify,bigdata*

***Abstract -** Internet is of life today. There are many applications that use the internet to access life, which one is SPOTIFY, Spotify is an online music platform that has many song lists and musicians in the world. However, on Spotify they have explicit songs, namely songs whose lyrics contain curse words, rude or discrimination. Big data is used for data analysis in the Snowflake application. Fortunately, we know that a new list of music and musicians has been created that uses explicit sentences by integrating data sourced from Kaggle in the form of CSV. After integration, we get the results of songs or musicians who use explicit words with the aim of knowing the content of the song.*

***Keywords:** Explicit,spotify,bigdata*

### **1. Pendahuluan**

Saat ini internet merupakan bagian yang tidak terpisahkan dalam kehidupan sehari-hari. Banyaknya aplikasi yang menggunakan internet untuk mengakses berbagai kebutuhan seperti sosial media[1], Selain kebutuhan internet, terdapat Big Data yang merupakan salah satu cara dalam menganalisis strategi untuk menemukan jawaban yang memiliki kemungkinan serta dapat mendukung dalam melakukan pengambilan sebuah keputusan[2].

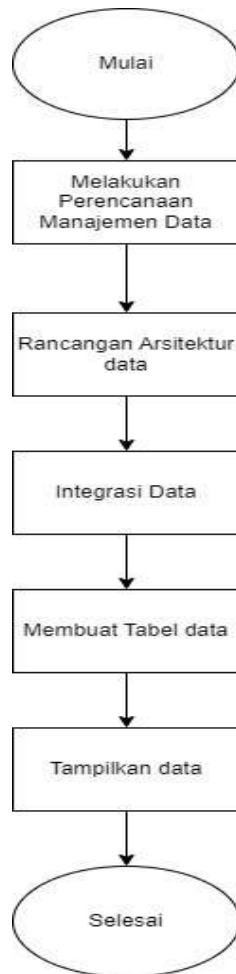
Sosial media memiliki perkembangan yang pesat saat ini yang menghasilkan budaya baru. Dan dapat diakses oleh seluruh lapisan Masyarakat untuk mencari berbagai informasi dan memiliki dampak terhadap Masyarakat[3]. Dengan adanya perkembangan internet, saat ini terdapat sosial media yang bergerak sebagai *online streaming* musik yaitu spotify. Spotify sendiri merupakan *platform* yang berasal dari Swedia yang digunakan dalam karya musik[4]. Banyaknya musisi yang berkarya melalui media spotify ini, terdapat karya karya memiliki catatan khusus yaitu lagu yang memiliki eksplisitas dengan lirik yang mengandung umpatan dan kritik yang keras. Saat ini spotify bisa diakses secara gratis ataupun berbayar dengan berbagai paket yang tersedia[5].

Untuk menganalisa data tersebut, penulis menggunakan perangkat lunak Snowflake dalam menganalisis dataset music pada spotify yang terdapat dari tahun 1950 – 2010. Sehingga dari data ini, penulis mendapatkan jenis jenis musik dari Musisi yang menggunakan kata-kata umpatan, kata kata kasar dan diskriminatif.

**2. Metode Penelitian**

Dalam penelitian ini, penulis merancang menggunakan metode kualitatif perangkat lunak Snowflake dan dataset yang diambil dari website keagle untuk list spotify. Dalam penelitian ini, penulis merancang tabel di Snowflake kemudian mengambil data – data yang diperlukan untuk di analisis sesuai dengan flowchart pada gambar 1.

**2.1. Alur Perancangan**



Gambar 1 Alur Perancangan

Berdasarkan alur penelitian, maka prosesnya adalah sebagai berikut :

1. Sebelum dilakukan analisa, maka diperlukan terlenih dahulu perencanaan manajemen data, untuk data yang diperlukan adalah ID,name,account\_name,artists,popularity dan explicit
2. Pada rancangan arsitektur data, format yang dibutuhkan adalah dalam bentuk csv

3. Setelah merancang arsitektur data, tabel yang sudah dibuat diintegrasikan dengan file spotify yang sudah diunduh dari website keagle.
4. Setelah diintegrasikan, maka selanjutnya dibuat tabel untuk menampilkan data hasil integrasi
5. Setelah itu data ditampilkan dengan hasil urutan Musisi pada spotify yang memiliki jumlah eksplisit terbanyak

### 3. Hasil dan Pembahasan

#### 3.1. Hasil Perancangan Tabel

Tabel data yang akan dianalisa dibuat dengan menampilkan ID, Account\_name, Judul, Artist, Popularity, dan Explicit dengan tipe data variable character dan jumlah karakter sebanyak 16777216. Jumlah karakter dimaksimalkan sesuai kapasitas maksimum ukuran karakter. Pada gambar 2 tabel sukses dibuat pada perangkat lunak snowflake.

Tabel 1. Perancangan tabel.

Nama	Tipe data	Jumlah Karakter
ID	Varchar	16777216
Account_name	Varchar	16777216
Judul	Varchar	16777216
Artists	Varchar	16777216
Popularity	Varchar	16777216
Explicit	Varchar	16777216

```

1050 Row(s) produced. Time Elapsed: 0.711s
yogaputr4#WH_SPOTIFY@DB_SPOTIFY.PUBLIC>create or replace table spotify_data (
    id VARCHAR,
    name VARCHAR,
    account_name STRING,
    artists STRING,
    popularity STRING,
    explicit STRING );
-----
| status
|-----
| Table SPOTIFY_DATA successfully created.
-----
    
```

Gambar 2. Tabel rancangan Basis Data.

#### 3.2. Hasil Integrasi Data Dengan SnowFlake

Hasil integrasi antara Snowflake yaitu snowSQL dengan data yang diunduh dari keagle tercantun pada gambar 3. Dari hasil tersebut data sudah berhasil terintegrasi file dalam bentuk csv dan snowSQL

```

1 Row(s) produced. Time Elapsed: 0.195s
yogaputr4#WH_SPOTIFY@DB_SPOTIFY.PUBLIC>copy into spotify_data
from @DB_SPOTIFY@PUBLIC.spotify_list
file_format = (
    type = csv
    skip_header = 1
    error_on_column_count_mismatch = false
    field_optionally_enclosed_by = '' );
-----
| file | status | rows_parsed | rows_loaded | error_limit | errors_seen | first_error | first_error_line | first_error_character | first_error_column_name |
-----
| spotify_list/spotify.csv.gz | LOADED | 1050 | 1050 | 1 | 0 | NULL | NULL | NULL | NULL |
-----
1 Row(s) produced. Time Elapsed: 1.673s
yogaputr4#WH_SPOTIFY@DB_SPOTIFY.PUBLIC>select * from SPOTIFY_DATA;
    
```

Gambar 3. Unggah file CSV

### 3.3. Hasil Analisis Data

Setelah dilakukan integrasi maka didapatkan hasil yaitu daftar lagu yang memiliki kata kata eksplisit yaitu lagu dengan judul *All Shock Up* dari Elvis Preley merupakan lagu yang memiliki kata kata yang mengandung umpatan atau kata kata kasar atau diskriminatif.

```

mysql> describe; Time elapsed: 1.090s
mysql> select * from SPOTIFY_DATA;
+-----+-----+-----+-----+-----+-----+
| ID | ACCOUNT_NAME | POPULARITY | EXPLICIT | JUDUL | ARTISTS |
+-----+-----+-----+-----+-----+-----+
| 1 | 3a23b11c8e19b3d1c4a0 | 117898 | 45 | All Shock Up | Elvis Presley |
| 2 | 221704 | 90 | 45 | I've Got You Under My Skin - Remastered 1998 | Frank Sinatra |
| 3 | 2c4c19a9e7e4f00f0e86 | 188933 | 40 | Snake Gets In Your Eyes | The Platters |
| 4 | 592842f30c8c808a10b | 347853 | 45 | What'd I Say, Pt. 1 & 2 | Ray Charles |
| 5 | 9ba32c0a8d7793a07e | 178048 | 40 | Dream A Little Dream of Me | Louis Armstrong |
| 6 | 908a9e83091250f1a2f | 135133 | 41 | That'll Be the Day | Buddy Holly |
| 7 | 8ba8e200b909e1a3e | 122893 | 47 | Don't Be Cruel | Elvis Presley |
| 8 | 4f011a1f40c19107f2c | 40 | Banana Boat (Day-O) | Harry Belafonte |
    
```

Gambar 4. Hasil Analisis Data Lagu Dengan Konten Eksplisit

Selain itu, didapatkan juga bahwa artist yang banyak mengandung konten eksplisit adalah album dari Fleetwood Mac yang memiliki total 648. Hasilnya terdapat pada gambar 5.

```

yogaputr4#WH_SPOTIFY@DB_SPOTIFY.PUBLIC>select ARTISTS, sum(implicit) as total
from "DB_SPOTIFY"."PUBLIC"."SPOTIFY_DATA"
group by ARTISTS
order by total desc;
+-----+-----+
| ARTISTS | TOTAL |
+-----+-----+
| Fleetwood Mac | 648 |
| Stevie Wonder | 601 |
| Elton John | 591 |
| Coldplay | 573 |
| Simon & Garfunkel | 532 |
| Adele | 502 |
| Whitney Houston | 490 |
| U2 | 474 |
| Taylor Swift | 451 |
| Prince | 451 |
| Billy Joel | 443 |
| Bruce Springsteen | 431 |
| The Beach Boys | 384 |
| Ed Sheeran | 371 |
| Amy Winehouse | 364 |
| Sam Cooke | 363 |
| The Beatles | 354 |
| Elvis Presley | 353 |
| Daryl Hall & John Oates | 336 |
| The Cure | 322 |
| Frank Sinatra | 311 |
| Ray Charles | 304 |
| Lana Del Rey | 303 |
| Sam Smith | 303 |
| The Platters | 301 |
| Justin Timberlake | 298 |
| Alicia Keys | 296 |
| ABBA | 295 |
| The Drifters | 291 |
| Phil Collins | 287 |
| Billie Holiday | 287 |
| The Everly Brothers | 285 |
    
```

Gambar 5. Hasil Analisis Data Album Dengan Konten Eksplisit.

### 4. Kesimpulan

Dari Hasil analisa data yang sudah dilakukan, dapat diambil kesimpulan bahwa dalam perkembangan internet kedepannya, perlu untuk mengetahui konten-konten pada sebuah aplikasi yang memiliki konten dengan catatan khusus seperti aplikasi *spotify* yang memiliki konten yang

mengandung unsur umpatan, kata kata kasar dan diskriminatif dan agar orang tua dapat mengawasi penggunaan internet kepada anak anak agar tidak mengakses konten tersebut.

### **Referensi**

- [1] Nur Amalia, K., Halim, U. (2022). "Penggunaan Internet sebagai Media Pembelajaran." In Jurnal Publish (Vol. 37, Issue 1).
- [2] Ferdiansyah, V., Irwan, M., Nasution, P. (2023). "Penerapan Teknologi Big Data Dalam Pengembangan Database Pendidikan. Jurnal Riset Manajemen", (3), 22–29. <https://doi.org/10.54066/jurma.v1i3.591>
- [3] Vinka, A. M., Michele, N., Industri, F. T., Rupa, S., Desain, D. (2021). "PENGARUH TEKNOLOGI INTERNET TERHADAP PENGETAHUAN MASYARAKAT JAKARTA SEPUTAR INFORMASI VAKSINASI COVID-19"(Vol. 8, Issue 1).
- [4] Andry, J. F., Tjee, C. (2019). "ANALISIS MINAT MAHASISWA MENDENGARKAN APLIKASI MUSIK BERBAYAR DAN UNDUHAN MUSIK GRATIS". JBASE - Journal of Business and Audit Information Systems, 2 (2). <https://doi.org/10.30813/jbase.v2i2.1727>
- [5] Febriani, A., Tiorida, E. (2019). "Pengaruh Perceived Value Spotify Premium Terhadap Kepuasan Pelanggan (Studi Pada Mahasiswa di Wilayah Bandung Raya)", IRWNS (Politeknik Negeri Bandung)
- [6] [www.kaggle.com](https://www.kaggle.com) (<https://www.kaggle.com/datasets/tamle507/spotify-all-out-playlists-songs-1950s-2010s?resource=download>)