

Sistem Rekomendasi Musik dengan Menggunakan Metode *K-Nearest Neighbor* (KNN)

Marcellino Rivaldo Pelaupessy^{a1}, I Ketut Gede Suhartana^{a2}

^aProgram Studi Teknik Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana Kampus UNUD Bukit Jimbaran, Bali, Indonesia
¹pel4upessy@email.com
²ikg.suhartana@unud.ac.id

Abstrak

Musik adalah hobi yang sudah menjadi kebiasaan sehari – hari kita, dimanapun dan kapanpun kita berada mengakses untuk mendengarkan music melalui aplikasi yang sudah tersedia di berbagai *platform*. Dengan itu sangatlah dibutuhkan sistem pemberi rekomendasi music untuk membantu pengguna aplikasi dapat mendengar musik sesuai dengan selera masing masing. *Metode K-Nearest Neighbor (KNN)* adalah algoritma yang populer digunakan dalam banyak penelitian. Tapi dengan penelitian ini dapat direncanakan akan juga menggunakan *String Matching* untuk mendapatkan hasil yang lebih optimal lagi.

Keywords: Musik, Sistem Rekomendasi, KNN, String Matching

1. Pendahuluan

Musik adalah kumpulan dari suara, irama, tempo, yang akan membuat sebuah rangkaian nada yang terdengar harmonis, pada awal abad pertengahan music diciptakan untuk acara Religi dan hal duniawi menggunakan alat music, dengan perkembangannya jaman music dapat dibuat oleh suara manusia saja, bahkan teknologi yang sudah berkembang.

Dengan adanya teknologi yang berkembang, pada tahun 1877 diciptakan sebuah mesin penyimpan lagu yang terbuat dari aluminium foil dan yang ditusuk oleh jarum yang bernama fonograf, perkembangan semakin pesat hingga membuat aluminium foil berubah menjadi piringan yang biasanya ditaruh ke dalam gramofon, hingga dengan teknologi berkembang lagu dapat bisa disimpan di dalam aplikasi dan tidak perlu repot untuk membuat piringan (Compact Disk)

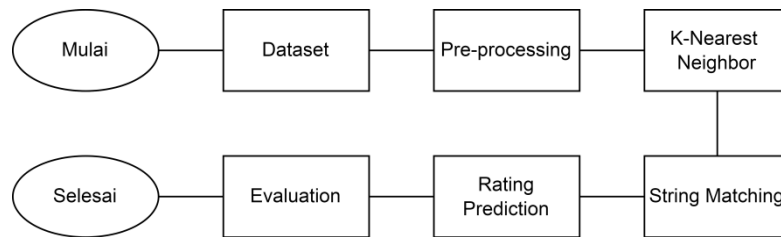
Teknologi dan Pengguna sedang sangat berkembang pesat, terutama pada bagian Musik yang sudah mempunyai akses yang luas dan gratis dimana saja. Untuk pengguna yang suka mendengarkan music melalui aplikasi yang tersedia, membutuhkan efisiensi untuk mencari music sesuai dengan selernya dari data yang sangat banyak jumlahnya, jadi dengan system pencarian rekomendasi music dapat membantu pengguna untuk lebih nyaman dengan aplikasinya sesuai dengan selera yang diinginkan

Salah satu rekomendasi metode yang digunakan agar lebih efisien ialah menggunakan *algoritma K-Nearest Neighbor (KNN)*. *K-Nearest Neighbor (KNN)* merupakan pengelompokan suatu data baru berdasarkan jarak data ke beberapa data. Metode ini dapat menyederhanakan algoritma perhitungan sehingga berpengaruh pada efisiensi waktu [1]. Penulis akan mencoba melakukan penelitian menggunakan metode *K-Nearest Neighbor(KNN)* guna untuk mendapatkan rekomendasi lagu agar lebih optimal

2. Metode Penelitian

2.1. Konsep Sebuah Sistem

Sebelum memulai penelitian, dibuat rancangan penelitian sebagai alur untuk memudahkan dalam melakukan penelitian



Gambar 1. Flowchart Konsep Sistem

2.2. Dataset

Dataset yang digunakan berasal dari *website* Kaggle, dimana dataset ini berformat file CSV. Dataset ini memiliki 232726 data lagu yang bersumber dari Spotify.

```
{ 'A Capella': 0, 'Alternative': 1, 'Anime': 2, 'Country': 3, 'Dance': 4, 'Electronic':
```

genre	artist_name	track_name	track_id	popularity	acousticness	d
0	Henri Salvador	C'est beau de faire un Show	0BRjO6ga9RKCKjfdqeFgWW	0	0.611	
1	Martin & les fées	Perdu d'avance (par Gad Elmaleh)	0BjC1NfoEOOusryehmNudP	1	0.246	
2	Joseph Williams	Don't Let Me Be Lonely Tonight	0CoSDzoNIKCRs124s9uTVy	3	0.952	
3	Henri Salvador	Dis-moi Monsieur Gordon Cooper	0Gc6TVm52BwZD07Ki6tIvf	0	0.703	
4	Fabien Nataf	Ouverture	0lusiXpMROhdEPvSI1FTQK	4	0.950	

Gambar 2. Dataset Kaggle

2.3. *K-Nearest Neighbor*

K-Nearest Neighbor adalah metode melakukan klasifikasi terhadap suatu object yang akan diklasifikasikan berdasarkan yang paling terdekat dengan object itu, Semua data juga bisa diklasifikasikan, berdasarkan teks, gambar, bahkan audio atau suara.

Rumus :

$$dis(x_1, x_2) = \sqrt{\sum_{i=0}^n (x_{1i} - x_{2i})^2}$$

$$dis = \sqrt{\sum_{i=0}^n (x_{1i} - x_{2i})^2 + (y_{1i} - y_{2i})^2 + \dots}$$

Gambar 3. KNN 1-dimension, KNN Euclidean distance

2.4. *String Matching*

String Matching adalah suatu algoritma pencocokan *string pattern* yang dibagi menjadi tiga yaitu Brute Force, Knuth-Morris-Pratt, dan Boyer-Moore, dan membandingkannya

menggunakan Regular Expression. Hasil dari pencarian string tergantung dari teknik pencocokannya. [2]

3. Hasil dan Pembahasan

3.1. Pre-processing data

Dataset yang sudah ada akan di-*label* untuk memudahkan kita pada nantinya memperhitungkan menggunakan *K-Nearest Neighbor*, karena pada nantinya data akan dibuat dalam bentuk matrix

```
{'00021ky6AyMbLP2tqij86e': 0, '000CzNKC8PEt1yC3L8dqW': 1, '000xQL6tZNLJzIrtIgxqS1': 2,
genre artist_name track_name track_id popularity acousticness danceability du
0 6 Henri Salvador C'est beau de faire un Show 817 0 0.611 0.389
1 6 Martin & les fées Perdu d'avance (par Gad Elmaleh) 832 1 0.246 0.590
2 6 Joseph Williams Don't Let Me Be Lonely Tonight 918 3 0.952 0.663
3 6 Henri Salvador Dis-moi Monsieur Gordon Cooper 1206 0 0.703 0.240
df_user = df.pivot(index="genre", columns="track_id", values="popularity").fillna(0)
df_user.head()
track_id 0 1 2 3 4 5 6 7 8 9 ... 35683 35684 35685
genre
0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 ... 0.0 0.0 0.0
1 0.0 0.0 0.0 0.0 49.0 0.0 0.0 0.0 0.0 55.0 ... 0.0 46.0 54.0
2 13.0 0.0 0.0 0.0 0.0 0.0 0.0 33.0 0.0 0.0 ... 0.0 0.0 0.0
3 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 ... 36.0 0.0 0.0
4 0.0 0.0 70.0 57.0 0.0 66.0 0.0 0.0 68.0 0.0 ... 0.0 0.0 0.0
5 rows x 35693 columns
```

Gambar 4. Pre-processing

3.2. K-Nearest Neighbor Process

Pada penelitian ini, implementasi penggunaan K-Nearest Neighbor dilakukan menggunakan algoritma bruteforce yang dimana mengambil sample matriks yang jaraknya paling kecil

3.3. String Matching

Penggunaan string matching pada penelitian kali ini ialah untuk mencari suatu pola yang satu dengan yang lain sehingga memudahkan kita menggunakan pendekatan melalui teks. String matching lebih tertuju pada satu pencarian atau lebih, semua kehadiran sebuah kata (lebih umum disebut pattern) dalam sebuah teks

3.4. Evaluasi

Pengukuran kinerja sistem yang telah dibangun sangat penting karena merupakan bentuk evaluasi kinerja algoritma yang telah dipilih, bagi peneliti yang ingin meningkatkan algoritma pembelajaran mesin pada domain masalah yang berbeda dan membandingkan hasil pengujian dengan yang lain untuk memahami metrik kinerja dan memilih kinerja yang paling tepat dari beberapa kemungkinan

Pada tahap ini akan dilakukan penghitungan system rekomendasi music dengan menggunakan sebuah nama track input yang akan dibandingkan dengan 4 track judul lagu rekomendasi dari

output dengan mengambil contoh judul lagu “Dis-moi Monsieur Gordon Cooper” yang akan dijadikan sebagai input, langkah ini bertujuan untuk mengetahui akurasi dari data tersebut [3]

```
recommender('Dis-moi Monsieur Gordon Cooper',mat_music,5)

Select the Movie : Dis-moi Monsieur Gordon Cooper Index : 3
Searching Recommendation for You...
3                               NaN
22                              Monsieur Boum Boum
24                              Quand je monte chez toi
11  The Hanging (Maverick - Original Motion Pictur...
14                              Keys of Love
Name: track_name, dtype: object
```

Gambar 5. Hasil Evaluasi

Dengan Hasil kita dapatkan melalui system yang kita buat, maka hasil close distance yang kita peroleh dari 4 judul yaitu

Track Lagu	Close Distance
Monsieur Boum Boum	0.89
Quand je monte chez toi	0.92
The Hanging (Maverick - Original Motion Pictures)	0.93
Keys of Love	0.95

Tabel 1. Hasil Close Distance

Dari hasil perhitungan yang sudah diperhitungkan menggunakan sistem, maka tingkat akurasi akan diperhitungkan secara normal, akurasi nilai yang sesuai dengan sebelumnya dibagi dengan jumlah data latih kemudian dikali dengan 100, maka keakurasian sistem yang dibuat adalah

$$\frac{3.69}{4} \times 100 = \%92.25$$

4. Kesimpulan

Berdasarkan hasil penelitian yang diperoleh dari menggunakan metode *K-Nearest Neighbor* pada dataset playlist dari aplikasi Spotify berjumlah 232726 track lagu dengan *close distance* dapat menjadi metode sistem rekomendasi yang baik bagi pengguna karena yang direkomendasikan sebagian besar serupa dengan item yang menjadi *input*.

References

- [1] R Yessivirna, Marji, D E Ratnawati. "Klasifikasi Suara Berdasarkan Gender (Jenis Kelamin) dengan Metode K-Nearest Neighbor (KNN)" 3-4.
- [2] Ernawati, A Johar, S Setiawan, "Implementasi Metode String Matching untuk Pencarian Portal Berita Berbasis Android (Studi Kasus : Harian Rakyat Bengkulu)"
- [3] F Shidiq, E W Hidayat, N I Kurniati, "Penerapan Metode K-Nearest Neighbor (KNN) Untuk Menentukan Ikan Cupang Dengan Ekstraksi Fitur Ciri Bentuk Dan Canny"