Aplikasi dengan Sistem Rekomendasi Genre Anime Menggunakan Metode KNN

**Achmad Rizky Zulkarnain1, Robert Antonius2**

Universitas Multi Data Palembang; Jl. Rajawali No.14, 9 Ilir, Kec. Ilir Tim. II, Kota Palembang, Sumatera Selatan 30113, (0711) 376400

Jurusan Ilmu Komputer dan Rekayasa, UMDP, Palembang

e-mail: [1rizkyzulkarnain@mhs.mdp.ac.id](mailto:rizkyzulkarnain@mhs.mdp.ac.id), [2robert.antonius@mhs.mdp.ac.id](mailto:robert.antonius@mhs.mdp.ac.id)

***Abstrak***

*[AbstrakMaksimal 200 kata berbahasa Indonesia dicetak miring dengan Times New Roman 11point. Abstrak harus jelas, deskriptif dan harus memberikan gambaran singkat masalah yang diteliti. Abstrak meliputi* ***alasan pemilihan topik atau pentingnya topik penelitian, metode penelitian dan ringkasan hasil.*** *Abstrak harus diakhiri dengan komentar tentang pentingnya hasil atau kesimpulan singkat. ]*

***Kata kunci***—klasifikasi, makanan, *convolutional neural network*

***Abstract***

*[A maximum 200 word abstract in English in italics with Times New Roman 11 point. Abstract should be clear, descriptive, and should provide a brief overview of the problem studied. Abstract topics include reasons for the selection or the importance of research topics, research methods and a summary of the results. Abstract should end with a comment about the importance of the results or conclusions brief. ]*

***Keywords***—*classification, food, convolutional neural network*

# 1. PENDAHULUAN

Di zaman Globalisasi ini, Anime sudah merambat ke seluruh dunia yang tadinya hanya di negara asalnya yaitu Jepang, tetapi sudah menyebar ke Seluruh Asia bahkan sudah memasuki Eropa. Dikarenakan film anime sudah terlalu banyak, genrenya juga sudah makin beraneka ragam seperti genre *isekai* yang baru memasuki dunia anime di tahun 2017. Oleh karena itu peniliti ingin membuat aplikasi yang memudahkan *audience* anime dalam mencari anime favoritnya dengan melihat anime apa saja yang sudah mereka tonton.

Adapun metode yang kami gunakan yaitu menggunakan metode KNN (*K-Nearest Neighboars)* dan dataset yang kami gunakan dari website *Kaggle.* Metode K-NN peneliti pakai dikarenakan dari jurnal [1], peneliti memakai metode K-NN dengan mengklasifikasi genre music didalamnya. Adapun *libary* yang kami pakai yaitu *sklearn* yang sudah disiapkan oleh *python.* Peneliti memakai bahasa *python* untuk pemogramannya dan *flutter* untuk *display-*nya. Tetapi peneliti bukan memakai proses klasifikasi tetapi agresi. Dikarenakan peneliti tidak mengelompokkan model yang kami dapatkan.

# 2. METODE PENELITIAN

Adapun data penelitian yang kami gunakan dari *public dataset.* Dataset yang kami gunakan dari *website kaggle*  adalah *Anime dataset Database*. Didalam dataset ini tidak *pure* sebuah dataset, dikarenakan atribut yang didapatkan di dalam dataset ini hanyalah pewarisan dari *website MAL* (*MyAnimeList).* Walaupun begitu, dataset ini cukup lengkap dengan mencakup 73,516 data user *MAL* dan 12,294 data anime didalmnya.

Adapun Atribut yang ada di dalam dataset tersebut :

|  |  |
| --- | --- |
| Anime.csv | |
| Anime-id | Uniqe id identifying anime |
| Name | Full name anime |
| Genre | Genre for anime |
| Type | Type anime (movie, OVA, etc) |
| Episodes | Episode of anime (1 if movie) |
| Rating | Average rating for anime |
| Member’s | Number of community members |

|  |  |
| --- | --- |
| Rating.csv | |
| User\_id | Non identifiable randomly generate user |
| Anime\_id | Anime that user rated |
| rating | Rating for that anime (-1 if user watched but not assign a rating) |

Adapun langkah-langkah yang kami lakukan sebelum melakukan training, yaitu merapikan fitur-fitur yang ada dan mengkompress file anime.csv. Berikut langkah-langkah yang kami lakukan :

1. Ada 2 data yang berisi rating user di dataset yaitu rating\_compelete dan anime\_list, peneliti memilih rating\_compelete dikarenakan data tersebut berisi isi dari anime\_list tetapi difilter hingga isinya hanya berisi anime yang sudah selesai ditonton saja.
2. Memfilter beberapa data yang mempunyai karakter jepang/asing yang meningkatkan berat file yang signifikan.
3. Memfilter beberapa data yang terkesan buruk, dengan mendrop beberapa column di file tersebut. Pertama mendrop column watched\_episodes dikarenakan peneliti tidak memperlukannya.
4. Memfilter rating yang bukan Currently Watching, Complete atau On Hold. Setelah ini peneliti mendrop column tersebut yang sudah di filter.
5. Setelah itu memfilter anime dengan tipe media tidak diketahui, skor yang tidak diketahui, skor yang dirating kurang dari 100 orang dikarenakan anime dibawah 100 rating sudah memiliki score unknown dan sudah difilter.
6. Memfilter anime dengan rating umur diatas 18 tahun
7. Memfilter orang yang menonton lebih dari 5000-10000 anime dikarenakan kemungkinan dia menge-set semua anime yang ada dengan complete dan itu sangat tidak wajar. 1000 anime itu “normal” bagi beberapa orang.
8. Memfilter orang yang menonton kurang dari 10 dikarenakan sulit dijadikan referensi.
9. Setelah difilter, peneliti mendrop semua yanhg sudah difilter tadi.

*2.1 Algoritma*

Nearest-Neighbors adalah algoritma untuk secara otomatis menghitung jarak antara 2 objek dari *libary* sklearn. Metric yang digunakan di set pada *cosine similiarity* karena kita mau melihat apakah anime A ditonton oleh *audience* yang sama dengan anime B. Algoritma yang peneliti pakai adalah brute sesuai rekomendasi *sklearn.* Peneliti menggunakan algoritma *brute* jika input datanya berbentuk *sparse matrix, metric precomputed,* dimensi jarak > 15, jumlah k lebih dari atau sama dengan setengah dari jumlah data, dan *effective metric* tidak dalam valid *metric*.

(1)

Keterangan :

= Sampel Data

= Sampel Data

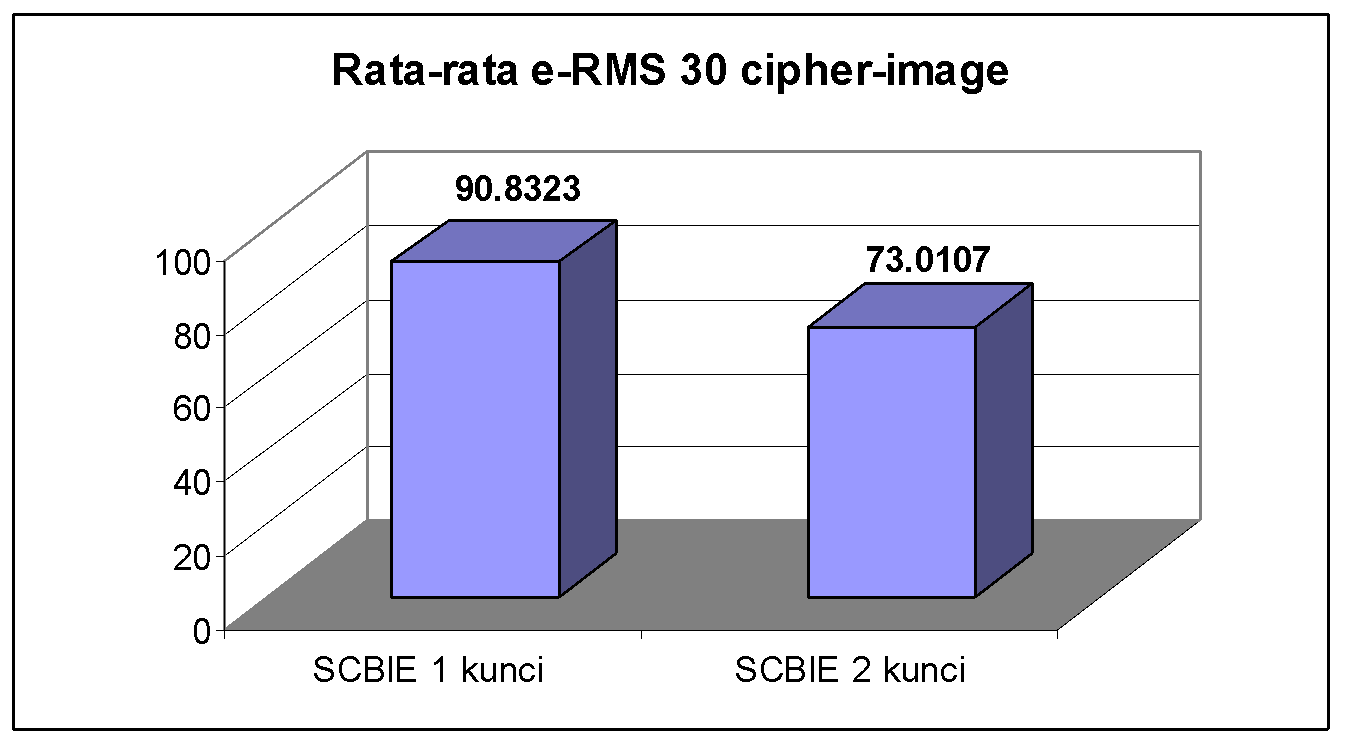
i = variabel data

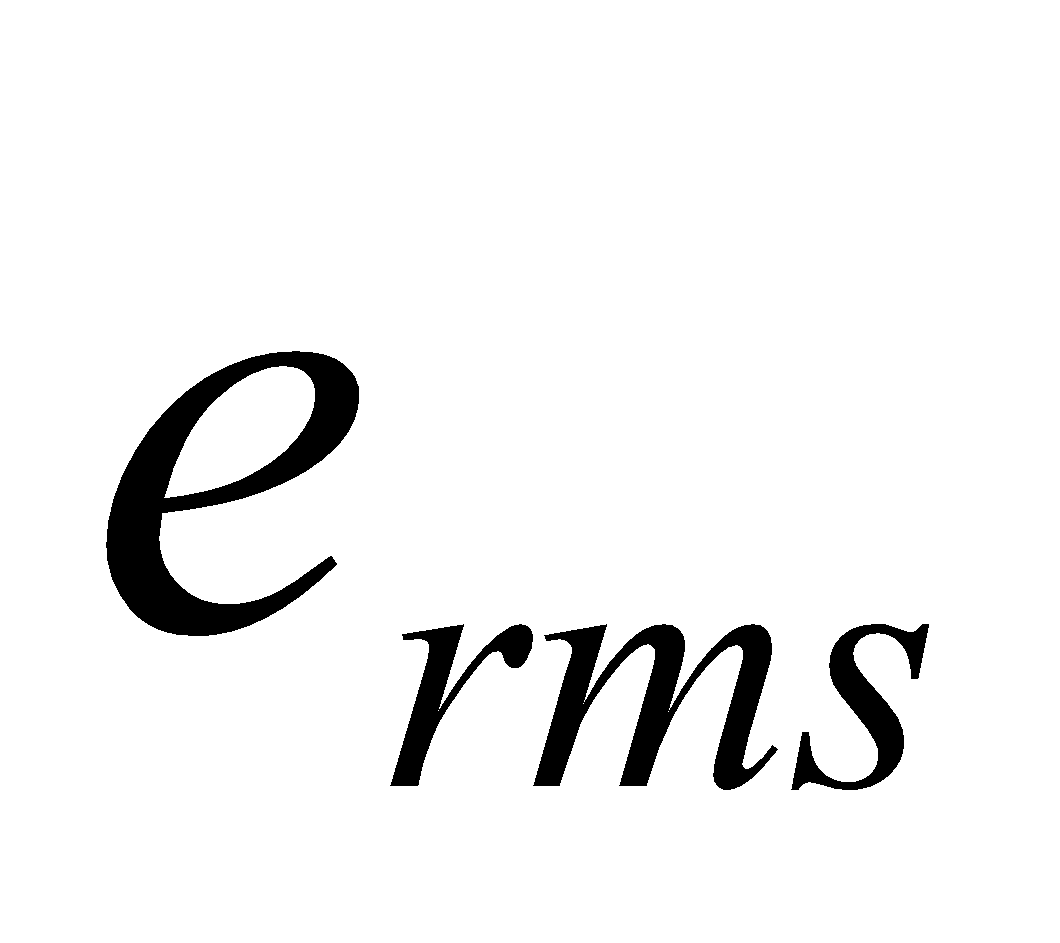
p = dimensi data

peneliti melakukan filter lagi yang mempunyai distance 1.0 dikarenakan anime seperti ini tidak memiliki tetangga, sehingga dapat di dikatakan bahwa data ini tidak bagus. Setelah itu, memuat semua list anime yang user sudah tonotn beserta skornya, dan difilter sehingga tidak memuat ID diatas 48492 karena dataset hanya memliki anime sampai ID tersebut.

# 3. HASIL DAN PEMBAHASAN

Pembahasan terhadap hasil penelitian dan pengujian yang diperoleh disajikan dalam bentuk uraian teoritik, baik secara kualitatif maupun kuantitatif.Hasil percobaan sebaiknya ditampilkan dalam berupa grafik atau pun tabel. Untuk grafik dapat mengikuti format untuk diagram dan gambar.



Gambar 2 Grafik perbandingan 

Tabel 1 Perbandingan Algoritma A dan Algoritma B

|  |  |  |  |
| --- | --- | --- | --- |
| Algoritma | Waktu Proses | Ketelitian | Memori |
| A | 120 ms | 98 % | 200 KB |
| B | 105 ms | 95 % | 415 KB |

# 4. KESIMPULAN

Kesimpulan harus mengindikasi secara jelas hasil-hasil yang diperoleh, kelebihan dan kekurangannya, serta kemungkinan pengembangan selanjutnya.

Kesimpulan dapat berupa paragraf, namun sebaiknya berbentuk point-point dengan menggunakan numbering atau bullet.

# 5. SARAN

Saran-saran untuk untuk penelitian lebih lanjut untuk menutup kekurangan penelitian. Tidak memuat saran-saran diluar untuk penelitian lanjut.

# UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada xxx yang telah memberi dukungan financial terhadap penelitian ini.

# DAFTAR PUSTAKA