

SISTEM REKOMENDASI FILM MENGGUNAKAN METODE *CONTENT-BASED FILTERING* STUDI KASUS MATERI DATA MINING DI SMK IDN BOARDING SCHOOL

Sri Lestari ¹, Muhammad Mufti Ramdhani ^{2*}

¹ Program Studi Sistem Informasi, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta, Indonesia.

^{2*} Program Studi Teknik Informatika, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta, Indonesia.

Email: sri.lestari1203@gmail.com ¹, mufti1453@gmail.com ^{2*}

Histori Artikel:

Dikirim 20 Juli 2023; *Diterima dalam bentuk revisi* 6 Agustus 2023; *Diterima* 15 Agustus 2023; *Diterbitkan* 10 September 2023. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

Abstrak

Sistem rekomendasi film telah menjadi topik penelitian yang menarik, karena dapat membantu pengguna dalam menemukan film yang sesuai dengan preferensi mereka. Metode Content-Based Filtering adalah salah satu metode yang umum digunakan dalam sistem rekomendasi film, yang berfokus pada konten film untuk membuat rekomendasi. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi film menggunakan metode Content-Based Filtering. Metode ini digunakan untuk melakukan rekomendasi film berdasarkan kemiripan fitur-fitur pada film yang telah ditonton oleh pengguna. Penelitian ini menggunakan studi kasus di SMK IDN Boarding School, dimana sistem rekomendasi film ini dibangun sebagai sarana pembelajaran materi data mining bagi siswa. Data film diambil dari API film dan diolah menggunakan algoritma untuk mengekstrak fitur dari konten film. Kemudian, sistem merekomendasikan film berdasarkan kesamaan fitur dengan film yang disukai oleh pengguna. Hasil pengujian menunjukkan bahwa sistem rekomendasi film yang dikembangkan memiliki kinerja yang baik dalam merekomendasikan film berdasarkan preferensi pengguna. Pengguna dapat memberikan umpan balik untuk meningkatkan akurasi rekomendasi dan sistem dapat terus ditingkatkan dengan menggunakan metode-metode yang lebih canggih.

Kata Kunci: Sistem Rekomendasi Film; Content-Based Filtering; Data Mining.

Abstract

Movie recommendation systems have become an interesting research topic, as they can assist users in finding movies that match their preferences. The Content-Based Filtering method is one of the methods commonly used in film recommendation systems, which focuses on film content to make recommendations. This study aims to develop a film recommendation system using the Content-Based Filtering method. This method is used to make film recommendations based on the similarity of the features of the films that have been watched by the user. This research uses case studies at SMK IDN Boarding School, where this film recommendation system was built as a means of learning data mining materials for students. Movie data is taken from the film API and processed using an algorithm to extract features from film content. Then, the system recommends films based on the similarity of features with the films that are liked by the user. The test results show that the developed film recommendation system has good performance in recommending films based on user preferences. Users can provide feedback to improve the accuracy of recommendations and the system can continue to be improved using more sophisticated methods.

Keyword: Movie Recommendation System; Content-Based Filtering; Data Mining.

1. Pendahuluan

Dalam era digital saat ini, jumlah film yang tersedia sangatlah banyak dan pengguna seringkali kesulitan dalam memilih film yang sesuai dengan preferensi mereka. Oleh karena itu, sistem rekomendasi film telah menjadi populer dalam beberapa tahun terakhir untuk membantu pengguna dalam menemukan film yang sesuai dengan selera mereka. Metode *Content-Based Filtering* adalah salah satu metode yang umum digunakan dalam sistem rekomendasi film. Metode ini memanfaatkan fitur-fitur dari konten film, seperti *genre*, aktor, sutradara, dan plot untuk merekomendasikan film yang memiliki kesamaan dengan film yang disukai oleh pengguna. Dalam konteks pembelajaran mesin, metode ini termasuk dalam kategori *supervised learning*.

SMK IDN *Boarding School* merupakan sebuah sekolah menengah kejuruan yang memiliki program studi rekayasa perangkat lunak. Materi data mining menjadi bagian dari kurikulum di program studi tersebut. Oleh karena itu, pengembangan sistem rekomendasi film menggunakan metode *Content-Based Filtering* dapat dijadikan sebagai sebuah studi kasus untuk mengaplikasikan materi data mining di SMK IDN *Boarding School*. Dalam penelitian ini, penulis akan mengimplementasikan sistem rekomendasi film menggunakan metode *Content-Based Filtering* pada materi data mining. Tujuan dari penelitian ini adalah untuk memberikan solusi bagi pengguna untuk menemukan film yang sesuai dengan selera mereka. Selain itu, penelitian ini juga diharapkan dapat menjadi kontribusi pada pengembangan kurikulum di SMK IDN *Boarding School*.

Penelitian dalam bidang sistem rekomendasi telah menjadi topik yang semakin menarik perhatian peneliti di berbagai bidang. Beberapa penelitian terkait dengan pengembangan sistem rekomendasi menggunakan metode *Content-Based Filtering* telah dilakukan. Fajriansyah, Adikara, dan Widodo (2021) melakukan penelitian tentang Sistem Rekomendasi Film menggunakan metode *Content-Based Filtering* [1]. Selain itu, Alkaff, Khatimi, dan Eriadi (2020) juga mengusulkan Sistem Rekomendasi Buku dengan menggunakan *Weighted Tree Similarity* dan *Content-Based Filtering* [2]. Badriyah, Fernando, dan Syarif (2018) menjelaskan Sistem Rekomendasi *Content-Based Filtering* dengan memanfaatkan algoritma Apriori [3]. Selain konten hiburan, penelitian juga mencakup rekomendasi produk. Larasati dan Februariyanti (2021) melakukan penelitian terkait Sistem Rekomendasi Produk Emina Cosmetics dengan metode *Content-Based Filtering* [4]. Kemudian, Harischandra, Pratama, Felix, dan Laia (2022) mengembangkan Aplikasi Pendukung Desain Interior yang menggunakan Sistem Rekomendasi Berdasarkan Nama Brand Perabot dengan metode *Content-Based Filtering* [5]. Safitri, Halilantar, dan Wahyuniar (2021) melakukan penelitian sejenis untuk Sistem Rekomendasi Skincare dengan metode *Content-Based Filtering* dan algoritma Apriori [6].

Selain itu, penelitian juga mencakup rekomendasi artikel ilmiah. Fitrianti, Fachurrozi, dan Yusliani (2018) menciptakan Sistem Rekomendasi Artikel Ilmiah Berbasis Web dengan menggunakan *Content-Based Learning* dan *Collaborative Filtering* [7]. Diah, Rahmad, dan Saleh (2019) mengimplementasikan Sistem Rekomendasi pada Pemilihan Kamera menggunakan metode *Content-Based* dan *Collaborative Filtering* [8]. Putri, Muchayan, dan Kamisutara (2022) fokus pada Sistem Rekomendasi Produk Pena Eksklusif dengan metode *Content-Based Filtering* dan TF-IDF [9]. Christian dan Kelvin (2022) mengembangkan Aplikasi Kursus Online Berbasis Web dengan Sistem Rekomendasi menggunakan metode *Content-Based Filtering* [10]. Huda, Fajarudin, dan Hadinegoro (2022) mengeksplorasi penggunaan TF-IDF Vector Similarity dalam Sistem Rekomendasi Artikel Berita berbasis *Content-Based Filtering* [11]. Selanjutnya, penelitian juga merambah ke bidang hiburan seperti film. Arfisko dan Wibowo (2022) mengembangkan Sistem Rekomendasi Film dengan pendekatan Hybrid Collaborative Filtering dan *Content-Based Filtering* [12]. Mondy, Wijayanto, dan Winarno (2019) berfokus pada Sistem Rekomendasi Kuliner dalam aplikasi Mangan dengan metode *Content-Based Filtering* [13]. Penelitian juga mencakup hiburan digital seperti video game. Pramestia dan Santiyasaa (2022) menerapkan metode *Content-Based Filtering* dalam Sistem Rekomendasi Video Game [14]. Terakhir, Wijaya dan Alfian (2018) mengembangkan Sistem Rekomendasi Laptop dengan pendekatan Collaborative Filtering dan *Content-Based Filtering* [15]. Selain itu, Liang dan Susanto (2022) membuat Aplikasi Rekomendasi Restoran dengan menggunakan metode K-Nearest Neighbors dan *Content-Based Filtering* [16].

Penelitian ini berfokus pada pengembangan sebuah Sistem Rekomendasi Film menggunakan metode *Content-Based Filtering*, dengan studi kasus yang dilakukan di SMK IDN Boarding School. Metode *Content-Based Filtering* adalah pendekatan yang digunakan untuk merekomendasikan item kepada pengguna berdasarkan karakteristik atau konten dari item yang telah mereka sukai atau interaksi sebelumnya. Selain itu, penelitian ini akan melibatkan siswa-siswa di SMK IDN Boarding School, mengingat pentingnya pemahaman mengenai teknologi seperti Data Mining dan sistem rekomendasi dalam dunia modern. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi yang berharga dalam pemahaman dan penerapan metode *Content-Based Filtering* dalam pengembangan Sistem Rekomendasi Film, serta menjadi dasar untuk pengembangan sistem serupa dalam konteks pendidikan dan hiburan di sekolah-sekolah lainnya. Dengan demikian, penelitian ini akan menggabungkan prinsip-prinsip Data Mining dengan metode *Content-Based Filtering* untuk menciptakan rekomendasi film yang lebih personal dan relevan bagi siswa SMK IDN Boarding School.

2. Metode Penelitian

Dalam membuat sebuah sistem rekomendasi film pada penelitian ini, dibutuhkan beberapa tahapan dan proses yang dilakukan, seperti analisis masalah, pengumpulan data, data preprocessing, penerapan filter, evaluasi dan implementasi GUI.

2.1 Analisis Masalah

Penelitian ini mengangkat tema mengenai sistem rekomendasi dengan menggunakan metode *Content-Based Filtering* dan menggunakan algoritma TF-IDF serta mengukur kemiripan tiap item dengan menggunakan metode cosine similarity. Pada proses tersebut, sistem rekomendasi diharapkan dapat memberikan referensi baru kepada pengguna mengenai berbagai macam film yang memiliki kemiripan dengan film yang diminati oleh para pengguna. Selain itu, permasalahan lainnya adalah bagaimana sistem dapat menampilkan rekomendasi film dan meminimalkan terjadinya kesalahan pada hasil rekomendasi.

2.2 Pengumpulan Data

Data yang digunakan pada sistem rekomendasi buku pada penelitian kali ini merupakan film yang didapatkan dari dataset private SMK IDN *Boarding School*. Jumlah data yang digunakan sebanyak 44555 film dan genre film yang diambil berasal dari beragam kategori mulai dari keluarga, komedi, aksi, fantasi, dan lain sebagainya.

2.3 Data Preprocessing

Pre-processing merupakan langkah awal yang dilakukan untuk mengubah data mentah menjadi informasi yang lebih bersih dan bisa diolah untuk proses selanjutnya. Pre-processing pada penelitian ini dilakukan dengan memanfaatkan Natural Language Processing (NLP) yang merupakan cabang dari kecerdasan buatan (artificial intelligence) yang berkaitan dengan kemampuan untuk memberi komputer pemahaman terhadap teks dan kata-kata yang diucapkan dengan cara yang sama seperti yang dapat dilakukan manusia. Pada proses ini, peneliti juga menggunakan Natural Language ToolKit (NLTK) yang merupakan tools pada Natural Language Processing (NLP) dengan menggunakan bahasa pemrograman *Python*. *Natural Language Toolkit* ini sangat mendukung proses pengolahan bahasa natural seperti *classification*, *tokenization*, *stemming*, *tagging*, *parsing*, dll. Untuk mengolah kata pada tahapan pre-processing data, peneliti menggunakan korpus yang merupakan kumpulan dari teks secara tertulis atau lisan yang tersimpan untuk kebutuhan suatu penyelidikan dan penelitian, salah satunya pada Natural Language Processing (NLP). Tahapan yang dilakukan untuk *preprocessing* data melalui beberapa tahapan seperti *case folding*, *tokenizing*, *filtering (stopword removal)*, dan *stemming*.

2.4 Penerapan Filter

Proses filter yang dilakukan adalah untuk memberikan bobot pada setiap term pada genre, durasi dan tahun rilis film.

2.5 Evaluasi

Evaluasi merupakan tahapan yang dilakukan untuk mengukur kinerja dari sistem yang telah dibuat. Data yang telah diolah dan telah melalui tahapan dari pre-processing hingga pengukuran kemiripan menggunakan cosine similarity perlu dievaluasi dan diukur seberapa akurat sistem tersebut. Evaluasi dilakukan dengan menggunakan precision sebagai ukuran tingkat ketepatan antara informasi yang diminta oleh pengguna dan jawaban yang diberikan oleh sistem. Dengan precision, akurasi dari sistem rekomendasi berbasis konten yang telah dibuat dapat diukur dan dievaluasi dengan baik.

2.6 Implementasi GUI

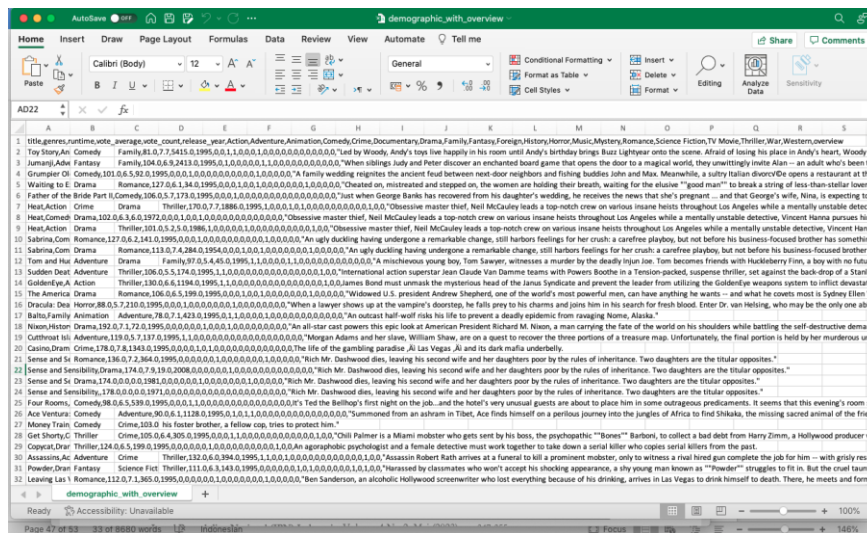
Sistem rekomendasi yang telah berhasil dibuat selanjutnya akan diimplementasikan dalam bentuk Graphical User Interface (GUI) pada sebuah website. Pembuatan website dilakukan dengan menggunakan Streamlit sebagai framework untuk sistem tersebut dan dijalankan secara local.

3. Hasil dan Pembahasan

Sistem rekomendasi film berhasil dikembangkan menggunakan metode Content-Based Filtering. Sistem ini mampu merekomendasikan film yang sesuai dengan preferensi pengguna berdasarkan kesamaan fitur dengan film yang disukai sebelumnya. Rekomendasi film didasarkan pada atribut-atribut seperti genre, durasi dan tahun rilis. Berikut hasil yang sudah dibuat.

3.1 Tampilan Dataset Film

Dataset berisi 44555 film yang disimpan dalam file berbentuk excel.



Gambar 1. Dataset Film

3.2 Tampilan Kode Program

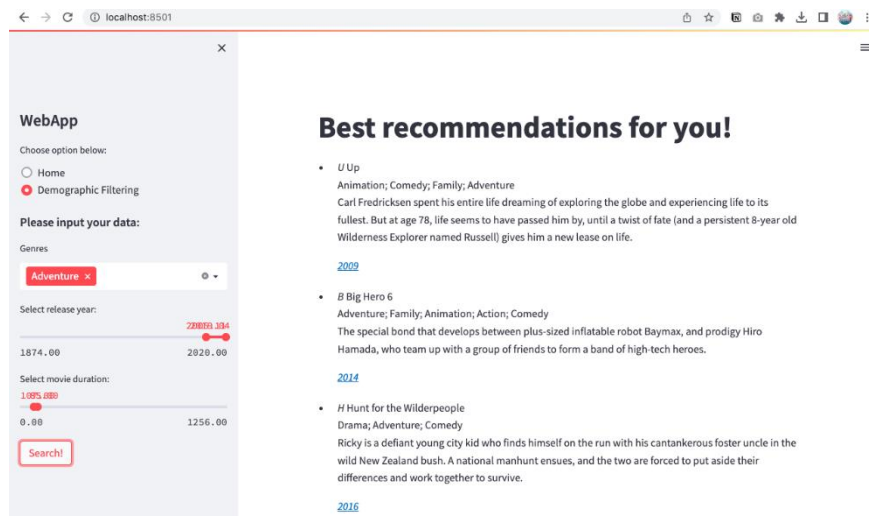
Kode program yang dibuat bertujuan untuk membaca dataset yang disimpan dalam bentuk excel, menerapkan filtering, weighting dan menampilkan data sesuai inputan user menggunakan framework streamlit.

```
1 import streamlit as st
2 import pandas as pd
3 from utils import filter_recommendation
4
5 def recommendation_system(mode):
6     df = pd.read_csv('./data/demographic_with_overview.csv')
7
8     if mode == 'Demographic Filtering':
9         st.title('Best recommendations for you!')
10        st.sidebar.subheader('Please input your data:')
11
12        # genre selection
13        genre = st.sidebar.multiselect('Genres', df.columns[6:-1])
14
15        # data for slider
16        runtime_min, runtime_q1, runtime_q3, runtime_max = df['runtime'].describe()
17        year_min, year_q1, year_q3, year_max = df['release_year'].describe()
18
19        # year slider
```

Gambar 2. Kode Program

3.3 Tampilan Implementasi GUI

Sistem rekomendasi yang telah berhasil dibuat selanjutnya akan diimplementasikan dalam bentuk *Graphical User Interface* (GUI) pada sebuah *website*. Pembuatan *website* dilakukan dengan menggunakan *streamlit* sebagai *framework* untuk sistem tersebut dan dijalankan secara local.



Gambar 3. Tampilan Web Sistem Rekomendasi Filem

4. Kesimpulan

Penelitian ini berhasil mengembangkan sistem rekomendasi film menggunakan metode *Content-Based Filtering* di lingkungan SMK IDN *Boarding School*. Dengan ekstraksi fitur dari konten film, seperti *genre*, durasi dan tahun rilis, sistem rekomendasi dapat memberikan rekomendasi film yang sesuai dengan preferensi pengguna berdasarkan kesamaan fitur dengan film yang disukai sebelumnya. Metode *Content-Based Filtering* memiliki kelebihan dalam memberikan rekomendasi film yang lebih spesifik dan personal kepada pengguna. Dengan tidak memerlukan kolaborasi antar pengguna, sistem dapat memberikan rekomendasi kepada pengguna yang baru dan mampu mengakomodasi preferensi individual dengan lebih baik. Meskipun demikian, metode *Content-Based Filtering* juga memiliki

keterbatasan dalam hal hanya mempertimbangkan atribut-atribut yang digunakan dalam ekstraksi fitur.

5. Ucapan Terima Kasih

Saya mengucapkan terima kasih kepada institusi pendidikan, SMK IDN *Boarding School*, yang telah memberikan dukungan dan kesempatan bagi saya untuk melakukan penelitian ini. Tanpa dukungan mereka, penelitian ini tidak akan terwujud. Saya juga mengucapkan terima kasih kepada semua rekan peneliti, keluarga, dan teman-teman yang memberikan dukungan, inspirasi, dan dorongan selama proses penelitian ini.

6. Daftar Pustaka

- [1] Fajriansyah, M., Adikara, P. P., & Widodo, A. W. (2021). Sistem Rekomendasi Film Menggunakan Content Based Filtering. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 5(6), 2188-2199.
- [2] Alkaff, M., Khatimi, H., & Eriadi, A. (2020). Sistem Rekomendasi Buku Menggunakan Weighted Tree Similarity dan Content Based Filtering. *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput*, 20(1), 193-202.
- [3] Badriyah, T., Fernando, R., & Syarif, I. (2018). Sistem Rekomendasi Content Based Filtering Menggunakan Algoritma Apriori. *Konferensi Nasional Sistem Informasi (KNSI) 2018*.
- [4] Larasati, F. B., & Februariyanti, H. (2021, Januari). SISTEM REKOMENDASI PRODUCT EMINA COSMETICS DENGAN MENGGUNAKAN METODE CONTENT - BASED FILTERING. *MISI (Jurnal Manajemen informatika & Sistem Informasi)*, 4(1), 45-54. DOI: <https://doi.org/10.36595/misi.v4i1.250>
- [5] Harischandra, A. R., Pratama, M. F. A., Felix, F., & Laia, A. P. (2022). Aplikasi Pendukung Desain Interior dengan Sistem Rekomendasi Berdasarkan Nama Brand Perabot Menggunakan Algoritma *Content-Based Filtering* Berbasis Web. *Jurnal SIFO Mikroskil*, 23(1), 1-16.
- [6] Safitri, D. A. N., Halilintar, R., & Wahyuniar, L. S. (2021, August). Sistem rekomendasi skincare menggunakan metode *Content-Based Filtering* dan algoritma apriori. In *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)* (Vol. 5, No. 2, pp. 242-248).
- [7] Fitrianti, B. S., Fachurrozi, M., & Yusliani, N. (2018). Sistem Rekomendasi Artikel Ilmiah Berbasis Web Menggunakan Content-based Learning dan Collaborative Filtering. *Generic Jurnal Ilmu Komputer dan Teknologi Informasi*, 10(1), 17-23.
- [8] Diah, I., Rahmad, I. F., & Saleh, A. (2019, April). IMPLEMENTASI RECOMMENDER SYSTEM PADA PEMILIHAN KAMERA MENGGUNAKAN CONTENT BASED DAN COLLABORATIVE FILTERING. *IT Journal*, 7(1), 1-12.
- [9] Putri, M. W., Muchayan, A., & Kamisutara, M. (2022). Sistem Rekomendasi Produk Pena Eksklusif Menggunakan Metode *Content-Based Filtering* dan TF-IDF. *JOINTECS (Journal of Information Technology and Computer Science)*, 7(1), 229-236.

- [10] Christian, Y., & Kelvin. (2022). RANCANG BANGUN APLIKASI KURSUS ONLINE BERBASIS WEB DENGAN SISTEM REKOMENDASI METODE *CONTENT-BASED FILTERING*. RABIT : Jurnal Teknologi dan Sistem Informasi Univrab, 7(1), 23-26.
- [11] Huda, A. A., Fajarudin, R., & Hadinegoro, A. (2022, Desember). Sistem Rekomendasi *Content-Based Filtering* Menggunakan TF-IDF Vector Similarity Untuk Rekomendasi Artikel Berita. Building of Informatics, Technology and Science (BITS), 4(3), 1679–1686.
- [12] Arfisko, H. H., & Wibowo, A. T. (2022, Juni). Sistem Rekomendasi Film Menggunakan Metode Hybrid Collaborative Filtering Dan *Content-Based Filtering*. e-Proceeding of Engineering, 9(3), 2149-2159.
- [13] Mondy, R. H., Wijayanto, A., & Winarno. (2019, Desember). RECOMMENDATION SYSTEM WITH *CONTENT-BASED FILTERING* METHOD FOR CULINARY TOURISM IN MANGAN APPLICATION. ITSMART: Jurnal Ilmiah Teknologi dan Informasi, 8(2), 65-72.
- [14] Pramestia, D. A., & Santiyasaa, I. W. (2022, November). Penerapan Metode *Content-Based Filtering* dalam Sistem Rekomendasi Video Game. Jurnal Nasional Teknologi Informasi dan Aplikasinya, 1(1), 229-234.
- [15] Wijaya, A. E., & Alfian, D. (2018). SISTEM REKOMENDASI LAPTOP MENGGUNAKAN COLLABORATIVE FILTERING DAN *CONTENT-BASED FILTERING* . Jurnal Computech & Bisnis, 12(1), 11-27.
- [16] Liang, S., & Susanto, J. (2022). Rancang Bangun Aplikasi Rekomendasi Restoran Menggunakan Metode K-Nearest Neighbors dan *Content-Based Filtering*. JURIKOM (*Jurnal Riset Komputer*), 9(1), 8-17. DOI: <http://dx.DOI.org/10.30865/jurikom.v9i1.3816>.