

# IMPLEMENTASI CRISP-DM PADA SISTEM REKOMENDASI BERBASIS KONTEN UNTUK PENENTUAN JENIS IKAN HIAS DALAM AKUARIUM

Oleh: Suko Tyas Pernanda<sup>1</sup>, Eri Eli Lavindi<sup>2</sup>, Amran Yobioktabera<sup>3</sup>, Aisyatul Karima<sup>4</sup>

<sup>1,2,3,4</sup>Jurusan Teknik Elektro, Politeknik Negeri Semarang

Jl. Prof. Sudarto, Tembalang, Semarang, 50275

E-mail: sukotyasp@polines.ac.id<sup>1</sup>, erielilavindi@polines.ac.id<sup>2</sup>, amranyobi@polines.ac.id<sup>3</sup>, aisyakarima@polines.ac.id<sup>4</sup>

## Abstrak

*Ikan hias air tawar terdiri dari banyak jenis yang biasa dipelihara oleh para penghobi. Penghobi pemula seringkali melakukan kesalahan dalam memilih jenis-jenis ikan yang dapat dicampur dalam sebuah aquarium sehingga jenis-jenis ikan yang dipelihara dalam aquarium yang sama tidak kompatibel antara satu jenis dengan lainnya. Hal ini dapat menyebabkan kematian dini pada ikan yang dipelihara. Sehingga diusulkan sebuah pendekatan berbasis data mining dan machine learning dalam menentukan jenis-jenis ikan hias yang dapat dipelihara secara bersamaan dalam sebuah akuarium. Penelitian ini mengikuti kerangka kerja CRISP-DM yang bertujuan untuk mengembangkan sistem rekomendasi ikan hias air tawar guna membantu para penghobi akuarium, terutama para pemula dalam memilih spesies-spesies ikan yang sesuai untuk akuarium mereka. Sistem ini menggunakan teknik web scraping untuk mengumpulkan data dari fishkeepingworld.com, sebuah platform daring yang berisi artikel tentang ikan hias air tawar. Data yang terkumpul diproses, ditokenisasi, dan diubah menjadi vektor kata menggunakan Word2Vec. Algoritma k-NN digunakan untuk menemukan spesies ikan serupa atau kompatibel. Akurasi sistem dievaluasi menggunakan Root Mean Square Error (RMSE) dengan menghitung rata-rata RMSE pada seluruh dataset. Pengujian dilakukan dengan memberikan beberapa nilai pada dimensi vektor untuk metode Word2Vec, jumlah K pada metode k-NN, dan metode perhitungan jarak pada k-NN. Sistem menghasilkan nilai RMSE terbaik dengan nilai 0.1329 pada K=3, jumlah vektor 400, dan perhitungan jarak dengan manhattan distance. Sistem rekomendasi yang dihasilkan diharapkan memberikan panduan yang bermanfaat bagi pengguna dalam meningkatkan pengalaman dalam hobi akuarium.*

**Kata Kunci:** CRISP-DM, ikan hias, sistem rekomendasi, k-NN.

## Abstract

*Freshwater ornamental fish consist of various species commonly kept by enthusiasts. Beginner hobbyists often make mistakes in selecting fish species that can coexist in the same aquarium, leading to compatibility issues and potential early mortality of the fish. Therefore, a data mining and machine learning-based approach is proposed to determine compatible ornamental fish species for simultaneous cohabitation in an aquarium. This research follows the CRISP-DM framework, which aims to develop a freshwater fish recommendation system to assist aquarium enthusiasts, especially beginners, in choosing suitable and compatible fish species for their aquariums. The system utilizes web scraping techniques to gather data from fishkeepingworld.com, an online platform featuring articles about exotic freshwater fish. The collected data is processed, tokenized, and transformed into word vectors using Word2Vec. A k-NN algorithm is employed to find similar or compatible fish species. The system's accuracy is evaluated using Root Mean Square Error (RMSE) by calculating the average RMSE across the entire dataset. Testing is conducted by providing various values for the Word2Vec vector dimensions, the number of neighbors (K) in the k-NN method, and the distance calculation method in k-NN. The system achieves the best RMSE value of 0.1329 at K=3, 400 vector dimensions, and using the Manhattan distance calculation method. The resulting recommendation system hopefully provides valuable guidance for users, enhancing their aquarium hobby experience.*

**Keywords:** CRISP-DM, ornamental fish, recommendation system, k-NN.

## 1. Pendahuluan

Akuarium tidak hanya merupakan hiasan visual yang memukau, tetapi juga menciptakan ekosistem yang rumit di dalamnya (Fellman, 2021). Keberhasilan sebuah akuarium tidak hanya bergantung pada estetika tata letak atau keindahan ikan yang dipilih, tetapi juga pada kompatibilitas antara spesies ikan yang ditempatkan bersama dalam sebuah akuarium (Golon, 2014).

Dalam dunia penghobi aquarium yang berkembang pesat, penting bagi setiap penghobi, terutama pemula, untuk mendapatkan panduan yang tepat dalam memilih ikan air tawar yang sesuai (Sharpe, 2021). Penghobi pemula sering memadukan beberapa jenis ikan hias dengan perilaku dan kebutuhan yang berbeda, dimana pada akhirnya dapat menyebabkan stres pada ikan, konflik, bahkan kematian (Sterling, 2022a).

Salah satu aspek yang tidak bisa diabaikan adalah kepatuhan terhadap *Animal Welfare Act 2006*. Dalam konteks undang-undang ini, perlakuan yang baik terhadap ikan dan keberlanjutan ekosistem akuarium menjadi fokus utama. Dengan merawat ikan sesuai dengan standar kesejahteraan hewan yang ditetapkan, para penghobi tidak hanya menciptakan lingkungan akuarium yang seimbang, tetapi juga mematuhi hukum yang mengatur perlindungan hewan (INJAF, 2017).

Penelitian ini diharapkan menjawab kebutuhan tersebut dengan menggunakan teknologi *data mining* dan *machine learning*. Penelitian ini diharapkan mampu memahami kompleksitas setiap spesies ikan, termasuk preferensi habitat, perilaku makan, dan temperamen, sehingga dapat mengetahui dinamika kelompok pada jenis-jenis ikan hias dan mengidentifikasi kemungkinan konflik atau keselarasan antar-spesies. Dengan memahami kebutuhan ikan hias dan menekuni hobi ini dengan kebijaksanaan ekologis, para penghobi tidak hanya mendapatkan kepuasan estetika, tetapi juga ikut berkontribusi pada keberlanjutan dan keseimbangan alam.

## 2. Tinjauan Pustaka

### 2.1 Sistem Rekomendasi

Sistem rekomendasi merupakan sebuah sistem yang menyediakan dan memberikan rekomendasi berdasarkan preferensi dari pengguna (Mondi et al., 2019; Rosita et al., 2022). Rekomendasi diberikan berdasarkan berbagai macam proses pengambilan keputusan seperti rekomendasi item yang sebaiknya dibeli berdasarkan pembelian sebelumnya, rekomendasi musik berdasarkan musik yang sering didengar, dan lain sebagainya (Mondi et al., 2019).

Macam-macam algoritma sistem rekomendasi juga digunakan untuk memberikan rekomendasi, diantaranya content-based filtering (Mondi et al., 2019; Tommy et al., 2019), apriori (Tommy et al., 2019), item-based (Rosita et al., 2022), collaborative (Gumantung Gusti et al., 2019), k-NN (Anamisa et al., 2023; Bayu Anwari & Yuliazmi, 2022; Farhan et al., 2023; Gumantung Gusti et al., 2019; Prasetya, 2017), atau dengan melakukan kombinasi dari beberapa metode tersebut (Tommy et al., 2019).

Pada kasus rekomendasi dimana data berupa teks, perlu dilakukan ekstraksi fitur untuk mendapatkan nilai vektor dari sebuah teks. Hal ini dapat dilakukan dengan metode TF-IDF (Hendrawan Rifky et al., 2022), atau Word Embedding seperti Word2Vec (Hendrawan Rifky et al., 2022; Nur Ghaniaviyanto Ramadhan, 2021; Sahbuddin & Agustian, 2022). Apabila dilakukan perbandingan antara kedua metode vektorisasi tersebut, maka Word2Vec memberikan hasil yang lebih baik daripada TF-IDF (Hendrawan Rifky et al., 2022; Ramadhanti & Mariyah, 2019). Hal ini dikarenakan metode Word2Vec dapat mengetahui semantik atau hubungan antar kata dalam suatu kalimat (Ramadhanti & Mariyah, 2019).

## 2.2 CRISP-DM

Penelitian ini menggunakan framework CRISP-DM dalam proses pelaksanaannya. CRISP-DM merupakan standar proses yang sering dipakai untuk proyek atau tugas tentang data mining (Khumaidi, 2020). Metode ini terdiri dari 6 tahapan, yaitu:

### a. *Business Understanding*

Tahap ini merupakan tahap paling awal dimana tujuan bisnis yang ingin dicapai melalui proyek data mining didefinisikan. Langkah yang dilakukan pada tahap ini diantaranya adalah mengidentifikasi masalah yang ingin dipecahkan dan memahami kebutuhan pengguna.

### b. *Data Understanding*

Tahap ini merupakan tahap pengumpulan data yang relevan untuk sebuah penelitian atau proyek. Data dapat dikumpulkan secara manual, atau melalui otomatisasi seperti web scraping, atau didapatkan dari sumber-sumber yang menyediakan open data, atau dari instansi yang menyediakan open data. Selanjutnya, dilakukan eksplorasi dan analisis data untuk memahami struktur, distribusi, dan karakteristik pada data.

### c. *Data Preparation*

Pada tahap ini dilakukan pengolahan data agar data dapat digunakan sebagai dasar untuk melakukan *modeling*. Umumnya dilakukan *data cleansing* dan *preprocessing* untuk mengatasi *missing values*, *outlier*, atau kesalahan lainnya. Selain itu, dilakukan pemilihan fitur agar sesuai dengan kebutuhan pemodelan.

### d. *Modeling*

Pada tahap ini dilakukan pemilihan teknik pemodelan yang sesuai (seperti regresi, klasifikasi, atau clustering) dengan menggunakan data yang telah dipersiapkan untuk melatih model. Selanjutnya, optimasi model dilakukan untuk memastikan akurasi dan kinerjanya.

### e. *Evaluation*

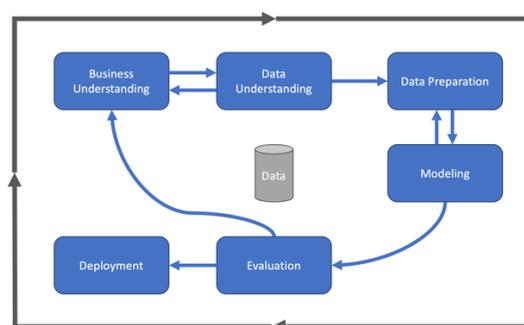
Pada tahap ini dilakukan pengukuran kinerja model menggunakan metrik evaluasi yang relevan (seperti akurasi,

presisi, atau RMSE). Tahap ini bertujuan untuk mengevaluasi apakah model memenuhi tujuan bisnis dan kebutuhan pengguna.

### f. *Deployment*

Tahap ini mengimplementasikan model ke dalam lingkungan produksi untuk memberikan hasil dan insight kepada pengguna atau pemangku kepentingan. Sehingga pengguna dapat memantau kinerja model yang diimplementasikan dan melakukan pembaruan jika diperlukan.

Gambaran tahapan pada CRISP-DM ditunjukkan pada Gambar 1.



Gambar 1. Diagram CRISP-DM

## 3. Metode Penelitian

Penelitian ini menggunakan CRISP-DM sebagai framework acuan dalam pelaksanaannya. Penelitian dibagi menjadi 6 tahap pelaksanaan sesuai dengan CRISP-DM Framework.

### 3.1 *Business Understanding*

Tahap pertama yang dilakukan adalah menganalisa dan memahami kebutuhan bisnis. Banyak pemula pada hobi ikan hias sering mencampuradukkan jenis-jenis ikan hias yang berbeda pada aquarium yang sama. Seringkali ikan hias yang dicampuradukkan tidak cocok antara satu jenis dengan jenis lainnya. Hal ini menyebabkan stress, hingga kematian dini pada ikan-ikan hias yang dipelihara (Fellman, 2021; Golon, 2014; Sharpe, 2021; Sterling, 2022b).

Pada penelitian ini, pengguna akan diminta satu jenis ikan yang ingin dipelihara sebagai masukan, kemudian mendapatkan luaran berupa beberapa jenis ikan yang sesuai untuk dipelihara bersamaan dengan

yang dipilih, sehingga meminimalisir pencampuradukkan jenis-jenis ikan hias yang tidak cocok satu dengan lainnya.

**3.2 Data Understanding**

Data diambil menggunakan teknik *web scraping* menggunakan library Selenium pada bahasa pemrograman python. Data diambil dari situs web <https://www.fishkeepingworld.com>. Data yang diambil ditunjukkan pada Gambar 2.



Gambar 2. Sumber data berupa website.

Data yang diambil berisi nama ikan hias dan deskripsi berupa teks yang mencakup kondisi lingkungan, ukuran, temperamen, dan lain-lain. Data tersebut kemudian disimpan dalam format *comma-separated value* (CSV). Hasil *scraping* dapat dilihat pada Gambar 3.

ID	title	content
0	endler	Endler's Livebearer is a colorful and attractive freshwater fish species from the Poeciliidae family.
1	silver-arowana	The silver arowana is a freshwater fish of the Osteoglossidae family. Native to South America, it is a popular and highly valued species.
2	frontosa-cichlid	The frontosa cichlid is a freshwater fish of the Cichlidae family. There are several different species, including the blue frontosa and the red frontosa.
3	red-zebra-cichlid	The red zebra cichlid is a freshwater fish of the Cichlidae family. These cichlids are found in the Great Lakes region of Africa.
4	dalmatian-molly	The dalmatian molly is a freshwater fish of the Poeciliidae family. This species is a popular choice for aquariums due to its unique black and white spots.
5	banjo-catfish	The banjo catfish is a freshwater fish of the Catfishidae family. The fish has an unusual banjo-shaped body and is known for its unique appearance.
6	peacock-eel	The peacock eel is a freshwater fish of the Mastacembelidae family. The species has a long, eel-like body in shades of brown, with eyespots decorating its tail and a yellow lateral stripe running from the fish snout to its caudal fin.
7	oranda-goldfish	The oranda goldfish is a freshwater fish of the Cyprinidae family. These fish are a popular choice for aquariums due to their unique, rounded bodies and colorful patterns.
8	cherry-barb	The cherry barb is a freshwater fish of the Cyprinidae family. The fish is small, bright red, and has a long, thin body.
9	female-betta-fish	Female betta fish are beautiful, rewarding fish for any aquarium keeper. Males are more aggressive and territorial than females.
10	beta-fish-care	Betta Fish Care: C First-time fish-keepers often underestimate how much care their new pets will need. Betta fish are sensitive to water fluctuations and prone to infections.
11	celestial-pearl-danio	The celestial pearl danio is a freshwater fish in the danio family. These fish have dark blue bodies with bright yellow spots and stripes.
12	blood-parrot-cichlid	The blood parrot cichlid is a freshwater fish of the cichlid family. Blood parrot cichlids are a unique species that are bred in captivity.
13	electric-blue-acara	The electric blue acara is a freshwater species of the cichlid family. The fish is related to the blue acara and is known for its vibrant blue color.

Gambar 3. Hasil pengambilan data.

**3.3 Data Preparation**

Pada tahap ini data mentah diolah sebelum digunakan untuk melatih model rekomendasi. Teks asli yang didapatkan dari proses *scraping* ditunjukkan pada Gambar 4.

The peacock eel is a freshwater fish of the Mastacembelidae family. The species has a long, eel-like body in shades of brown, with eyespots decorating its tail and a yellow lateral stripe running from the fish snout to its caudal fin. While peaceful, this species isn't suitable for beginner aquarists because the fish are sensitive to water fluctuations and prone to infections. Peacock Eel Facts & Overview Appearance & Behavior Peacock Eel Tank Requirements Common Peacock Eel Health Issues and Diseases Peacock Eel Tank Mates Peacock Eel Diet and Feeding Breeding the Freshwater Peacock Eel.

Gambar 4. Teks Hasil Scraping

Langkah-langkah yang dilakukan pada tahap ini yaitu:

- Konversi Teks ke Huruf Kecil, yaitu tahap dimana semua teks diubah menjadi huruf kecil untuk memastikan konsistensi dan menghindari perbedaan huruf besar dan kecil dalam analisis teks.
- Tokenisasi, yaitu pembagian teks deskripsi menjadi token atau kata-kata terpisah. Hal ini dilakukan untuk membantu memahami kata demi kata dalam sebuah teks.
- Lemmatization*, yaitu proses yang dilakukan untuk mengonversi kata-kata dalam teks ke bentuk dasarnya (kata dasar) untuk memperjelas makna. Contohnya, kata-kata seperti "berenang" akan diubah menjadi "renang".
- Stop Words Removal*: Selain itu, kata-kata umum (stop words) seperti "dan", "atau", dan "yang" juga dihilangkan karena kata-kata tersebut berulang sangat sering di banyak kalimat. Sehingga kata-kata tersebut tidak memberikan kontribusi signifikan dalam pemahaman semantik.
- Symbol Cleansing*, yaitu tahap untuk menghilangkan tanda baca seperti koma dan titik. Tanda baca tersebut dihilangkan karena tidak memiliki makna semantik.

Setelah melalui langkah-langkah tersebut, teks deskripsi telah diproses menjadi bentuk yang lebih bersih dan seragam seperti ditunjukkan pada Gambar 5.

```
peacock eel freshwater fish mastacembelidae family specie long
eel-like body shade brown eyespot decorating tail yellow lateral
stripe running fish snout caudal fin peaceful specie suitable
beginner aquarists fish sensitive water fluctuation prone infection
table content peacock eel fact overview appearance behavior
peacock eel tank requirement common peacock eel health issue
disease peacock eel tank mate peacock eel diet feeding breeding
freshwater peacock eel get peacock eel aquarium peacock eel faq
peacock eel fact overview scientific name macrogathus siamensis
common name peacock eel
```

Gambar 5. Hasil Preprocessing

Selanjutnya, dilakukan *feature extraction* pada teks yang telah dilakukan *preprocessing*. *Feature extraction* dilakukan menggunakan algoritma word2vec untuk mengetahui nilai semantik pada masing-masing kata. Luaran dari tahap ini adalah nilai vektor dari masing-masing kata. Hasil proses ini ditunjukkan pada Gambar 6.

```
[-0.01360063 0.45024241 -0.21452959 -0.03750757 0.21682188 -0.62697646
0.02001297 0.76837435 -0.39432867 -0.43001667 -0.1510052 -0.43737817
0.09073113 0.26149742 0.43552468 -0.37749959 -0.16948379 -0.53437075
-0.12095037 -0.69258528 0.53638966 0.05349278 0.24224526 -0.09917243
-0.18045022 -0.40422263 -0.07423973 -0.37472388 -0.44001766 0.08342592
0.28255023 0.32269401 0.23601753 -0.56166382 -0.09619438 0.35876464
-0.06421623 -0.18783122 -0.31877249 -0.90931819 0.15870919 -0.5138987
-0.28047702 0.45239747 0.6773256 0.06543603 -0.68319871 0.0542906
0.6242186 -0.10884207 0.41306436 -0.15894011 0.51903348 -0.01295026
-0.02501534 0.19062069 0.19963586 -0.03138964 -0.62493591 0.25331507
0.11003498 0.11462481 -0.50783972 -0.08761354 -0.37729025 0.51470314
0.15228466 0.16666517 -0.85293829 0.49132975 -0.30866255 0.07303166
0.52232817 -0.30479124 0.51641073 0.49915366 -0.02964419 0.01328135
-0.34053849 0.18780826 -0.27899378 0.01224678 -0.2185605 0.36144483
-0.07698939 -0.15166368 0.41521377 0.21644665 0.76061601 0.04325597
0.23406546 0.02192654 0.55244859 0.02843697 0.07140143 0.35813424
0.05611131 -0.32636906 0.1522657 -0.10105887]
```

Gambar 6. Hasil Feature Extraction.

### 3.4 Modeling

Pada tahap ini dilakukan permodelan menggunakan metode k-NN. Langkah-langkah yang dilakukan pada algoritma k-NN dalam penelitian ini adalah sebagai berikut:

- Pengukuran Kemiripan: Setiap deskripsi ikan dipetakan ke dalam ruang vektor. Kemudian, untuk mencari ikan terdekat dengan ikan input, k-NN menggunakan metrik cosine similarity dan manhattan distance.
- Menentukan Tetangga Terdekat: Algoritma k-NN mencari k tetangga terdekat dari ikan input. Nilai k menunjukkan jumlah tetangga yang akan digunakan untuk memberikan rekomendasi. Semakin besar nilai k, semakin banyak tetangga yang

dipertimbangkan.

- Memberikan Rekomendasi: Setelah tetangga terdekat ditentukan, ikan-ikan ini direkomendasikan kepada pengguna sebagai ikan yang mungkin cocok berdasarkan kemiripan konten deskripsi. Pengguna akan menerima daftar ikan terdekat berdasarkan kesamaan konten deskripsi mereka dengan ikan input.

### 3.5 Evaluation

Dalam penelitian ini, proses evaluasi dilakukan menggunakan metrik Root Mean Square Error (RMSE). RMSE mengukur seberapa akurat sistem dalam memprediksi preferensi pengguna terhadap ikan air tawar yang akan direkomendasikan. Semakin rendah nilai RMSE, semakin baik kualitas rekomendasi yang diberikan oleh sistem.

Beberapa parameter yang digunakan untuk proses pengujian yaitu jumlah K (3, 5, 10), metode perhitungan jarak (cosine similarity, manhattan distance), dan jumlah dimensi vektor pada metode word2vec (100, 200, 400).

## 4. Hasil dan Pembahasan

Pengujian dilakukan dengan mengkombinasikan 3 parameter yang telah disebutkan sebelumnya yaitu nilai K, metode perhitungan jarak, dan jumlah dimensi pada Word2Vec. Nilai K yang diujikan yaitu 3, 5, dan 10. Metode perhitungan jarak yang diujikan adalah cosine similarity dan manhattan distance. Sedangkan jumlah dimensi vektor yang diujikan adalah 100, 200, dan 400.

Pengujian dengan nilai K=3 mendapatkan hasil seperti pada Tabel 1, dimana nilai RMSE paling kecil didapatkan dengan kombinasi dimensi vektor sejumlah 400 dan perhitungan jarak menggunakan manhattan distance dengan nilai RMSE sebesar 0.1329.

Tabel 1. Hasil Pengujian Model K=3

Dimensi Vektor	Perhitungan Jarak	RMSE
100	cosine	0.2019
100	manhattan	0.1647
200	cosine	0.1974

200	manhattan	0.1471
400	cosine	0.1829
400	manhattan	<b>0.1329</b>

Pengujian dengan nilai  $K=5$  mendapatkan hasil seperti pada Tabel 2, dimana nilai RMSE paling kecil didapatkan dengan kombinasi dimensi vektor sejumlah 400 dan perhitungan jarak menggunakan mahattan distance dengan nilai RMSE sebesar 0.1632.

**Tabel 2. Hasil Pengujian Model  $K=5$**

Dimensi Vektor	Perhitungan Jarak	RMSE
100	cosine	0.2477
100	manhattan	0.2005
200	cosine	0.2332
200	manhattan	0.1792
400	cosine	0.2272
400	manhattan	<b>0.1632</b>

Pengujian dengan nilai  $K=10$  mendapatkan hasil seperti pada Tabel 3, dimana nilai RMSE paling kecil didapatkan dengan kombinasi dimensi vektor sejumlah 400 dan perhitungan jarak menggunakan mahattan distance dengan nilai RMSE sebesar 0.2020.

**Tabel 3. Hasil Pengujian Model  $K=10$**

Dimensi Vektor	Perhitungan Jarak	RMSE
100	cosine	0.2926
100	manhattan	0.2401
200	cosine	0.2746
200	manhattan	0.2183
400	cosine	0.2593
400	manhattan	<b>0.2020</b>

## 5. Kesimpulan dan Saran

### 5.1. Kesimpulan

Sistem rekomendasi berbasis konten dapat diterapkan pada kasus kecocokan pemilihan jenis-jenis ikan hias air tawar dalam satu akuarium dengan rata-rata nilai error yang kecil. Sehingga dapat digunakan sebagai acuan untuk para pemelihara ikan hias dalam memilih jenis-jenis ikan hias yang sesuai dalam akuariumnya.

Berdasarkan hasil pengujian di atas, dapat disimpulkan bahwa semakin kecil jumlah  $K$ , maka semakin kecil nilai

errornya. Hal ini disebabkan karena semakin sedikit pula jumlah  $K$  yang dijadikan pertimbangan dalam menentukan nilai error. Selanjutnya, dapat disimpulkan pula bahwa semakin besar dimensi vektor pada metode Word2Vec, akan menghasilkan nilai error semakin kecil. Hal ini dapat disebabkan karena semakin besar dimensi vektor, kemungkinan untuk mendapatkan fitur yang unik pada setiap kata akan semakin besar. Selanjutnya, perhitungan jarak menggunakan *manhattan distance* menghasilkan nilai error yang lebih kecil dibandingkan dengan cosine similarity. Hal ini dapat disebabkan karena dimensi yang digunakan pada perhitungan *manhattan distance* adalah vektor dengan dimensi yang digunakan dalam penelitian ini relatif tinggi yaitu 100, 200, dan 400.

Dari seluruh kombinasi parameter tersebut, hasil terbaik didapatkan dengan memberikan parameter berupa  $K=4$ , dimensi Word2Vec sejumlah 400, dan perhitungan jarak menggunakan manhattan distance. Kombinasi tersebut menghasil nilai RMSE paling kecil yaitu sebesar **0.1329**.

### 5.2. Saran

Penelitian ini masih dapat dikembangkan dengan menyimpan data pilihan user yang disimpan untuk digunakan sebagai acuan sistem rekomendasi selanjutnya. Dengan menggunakan data pilihan user yang disimpan tersebut, dapat dilakukan pengembangan tahapan *modeling* dengan menggunakan metode sistem rekomendasi lainnya seperti *collaborative filtering* yang dapat mengatasi keterbatasan dari metode *content-based* yang digunakan pada penelitian ini.

## DAFTAR PUSTAKA

Anamisa, D. R., Jauhari, A., & Ayu Mufarroha, F. (2023). K-Nearest Neighbors Method for Recommendation System in Bangkalan’s Tourism. *ComTech: Computer, Mathematics and Engineering Applications*, 14(1), 33–44.

- <https://doi.org/10.21512/comtech.v14i1.7993>
- Bayu Anwari, V., & Yuliazmi. (2022). Implementasi Algoritma K-Nearest Neighbors Pada Analisis Sentimen Masyarakat Terhadap Penerapan Pemberlakuan Pembatasan Kegiatan Masyarakat. *Sistem Komputer Dan Teknik Informatika*, 5(1), 72–81. <https://doi.org/10.36080/skanika.v5i1.2912>
- Farhan, M., Purbolaksono, M. D., & Astuti, W. (2023). Sentiment Analysis of Practo App Reviews using KNN and Word2Vec. *Building of Informatics, Technology and Science (BITS)*, 5(1). <https://doi.org/10.47065/bits.v5i1.3598>
- Fellman, S. (2021). *THE AQUARIUM AS AN ECOSYSTEM: AN OLD IDEA THAT SOMEHOW ALWAYS SEEMS NEW.* <https://tanninaquatics.com/blogs/the-tint-1/the-aquarium-as-an-ecosystem-an-old-idea-that-somehow-always-seems-new>
- Golon, C. (2014). *Tips for Freshwater Aquarium Success.* <https://www.vetstreet.com/our-pet-experts/tips-for-freshwater-aquarium-success>
- Gumantung Gusti, I., Nasrun, M., Astuti Nugrahaeni, R., & Komputer, T. (2019). REKOMENDASI SISTEM PEMILIHAN MOBIL MENGGUNAKAN K-NEAREST NEIGHBOR (KNN) COLLABORATIVE FILTERING. In *Jurnal TEKTRIKA* (Vol. 4, Issue 1). <https://doi.org/10.25124/tektrika.v4i1.1846>
- Hendrawan Rifky, I., Utami, E., & Hartanto Dwi, A. (2022). Analisis Perbandingan Metode Tf-Idf dan Word2vec pada Klasifikasi Teks Sentimen Masyarakat Terhadap Produk Lokal di Indonesia. *Smart Comp: Jurnalnya Orang Pintar Komputer*, 11(3). <https://doi.org/10.30591/smartcomp.v11i3.3902>
- INJAF. (2017). *Raising your concerns.* <https://injaf.org/the-think-tank/raising-your-concerns/>
- Khumaidi, A. (2020). DATA MINING FOR PREDICTING THE AMOUNT OF COFFEE PRODUCTION USING CRISP-DM METHOD. *Jurnal Techno Nusa Mandiri*, 17(1), 1–8. <https://doi.org/10.33480/techno.v17i1.1240>
- Mondi, R. H., Wijayanto, A., & Winarno. (2019). RECOMMENDATION SYSTEM WITH CONTENT-BASED FILTERING METHOD FOR CULINARY TOURISM IN MANGAN APPLICATION. *ITS SMART: Jurnal Ilmiah Teknologi Dan Informasi*, 8(2). <https://doi.org/https://doi.org/10.20961/itsmart.v8i2.35008>
- Nur Ghaniaviyanto Ramadhan. (2021). Indonesian Online News Topics Classification using Word2Vec and K-Nearest Neighbor. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(6), 1083–1089. <https://doi.org/10.29207/resti.v5i6.3547>
- Prasetya, C. S. D. (2017). Sistem Rekomendasi Pada E-Commerce Menggunakan K-Nearest Neighbor. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 4(3), 194. <https://doi.org/10.25126/jtiik.201743392>
- Ramadhanti, N. R., & Mariyah, S. (2019). Document Similarity Detection Using Indonesian Language Word2vec Model. *2019 3rd International Conference on Informatics and Computational Sciences (ICICoS)*, 1–6. <https://doi.org/10.1109/ICICoS48119.2019.8982432>
- Rosita, A., Puspitasari, N., & Kamila, V. Z. (2022). REKOMENDASI BUKU PERPUSTAKAAN KAMPUS

- DENGAN METODE ITEM-BASED COLLABORATIVE FILTERING. *Sebatik*, 26(1), 340–346. <https://doi.org/10.46984/sebatik.v26i1.1551>
- Sahbuddin, M., & Agustian, S. (2022). Support Vector Machine Method with Word2vec for Covid-19 Vaccine Sentiment Classification on Twitter. *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, 6(1), 288–297. <https://doi.org/10.31289/jite.v6i1.7534>
- Sharpe, S. (2021). *Common Mistakes When Starting a New Aquarium*. <https://www.thesprucepets.com/common-new-aquarium-mistakes-1380712>
- Sterling, I. (2022a). *Why Do My Fish Keep Dying? 11 Reasons & How to Prevent It*. <https://fishlab.com/why-your-fish-are-dying/>
- Sterling, I. (2022b). *Why Do My Fish Keep Dying? 11 Reasons & How to Prevent It*. <https://fishlab.com/why-your-fish-are-dying/>
- Tommy, L., Kirana, C., & Lindawati, V. (2019). RECOMMENDER SYSTEM DENGAN KOMBINASI APRIORI DAN CONTENT-BASED FILTERING PADA APLIKASI PEMESANAN PRODUK. *Jurnal Teknoinfo*, 13(2), 84. <https://doi.org/10.33365/jti.v13i2.299>

