

Implementasi Algoritma Machine Learning untuk Rekomendasi Produk pada Sistem Penjualan Berbasis Web Menggunakan Dataset Publik

Yuri Tri Wibowo^{1*}, Nibras Faiq Muhammad², Pramono³

¹Teknik Informatika/Fakultas Ilmu Komputer

Universitas Duta Bangsa

^{1*}220103257@mhs.udb.ac.id

²Teknik Informatika/Fakultas Ilmu Komputer

Universitas Duta Bangsa

²nibras_faiqmuhammad@udb.ac.id

³Teknik Informatika/Fakultas Ilmu Komputer

Universitas Duta Bangsa

³pramono@udb.ac.id

Abstrak— Perkembangan teknologi informasi mendorong pemanfaatan sistem rekomendasi untuk membantu pengguna menemukan produk yang sesuai dengan preferensinya. Pada sistem penjualan berbasis web, pengguna sering mengalami kesulitan dalam memilih produk karena banyaknya pilihan yang tersedia. Penelitian ini bertujuan mengimplementasikan Machine Learning pada sistem sistem penjualan berbasis web untuk menghasilkan rekomendasi produk menggunakan dataset publik. Dataset yang digunakan berasal dari Book Recommendation data yang diperoleh dari platform Kaggle dan memanfaatkan data rating pengguna sebagai dasar rekomendasi. Metode yang di terapkan adalah Collaborative Filtering untuk menghitung tingkat kemiripan antar pengguna menggunakan Cosine Similarity dan Algoritma K-Nearest Neighbor untuk menentukan pengguna dengan preferensi yang paling mirip. Sistem dikembangkan menggunakan Framework Laravel sebagai platform web, MySQL sebagai basis data ,serta Phyton untuk proses perhitungan rekomendasi. Hasil penelitian menunjukkan bahwa sistem mampu mengolah data rating pengguna dan menghasilkan rekomendasi buku yang relevan berdasarkan pola preferensi pengguna lain yang memiliki karakteristik serupa. Implementasi metode Collaborative Filtering dan K-Nearest Neighbor pada sistem penjualan berbasis web dapat membantu pengguna memperoleh rekomendasi yang lebih personal sehingga meningkatkan pengalaman pengguna dalam proses pencarian produk.

Kata Kunci—Machine Learning, Sistem Rekomendasi, Collaborative Filtering, K-Nearest Neighbor, Laravel

Abstract— The rapid Development of information technology has encouraged the implementation of recommendation systems to assist user in finding products that match their preferences, in web-based sales systems, users often face difficulties in selecting products due to the large number of available options. This study aims to implement Machine Learning in a web-based sales system to generate product recommendations using a public dataset. The dataset used in this research is the Book Recommendation Dataset obtained from Kaggle, which utilizes user rating data as the basis for recommendations. The proposed method combines Collaborative Filtering to calculate user similarity using Cosine Similarity and the K-Nearest Neighbor algorithm to identify users with the most similar preferences. The System was developed using the Laravel Framework as the web platform, MySQL as the database management system, and Phyton for recommendation processing. The results indicate that the system is cable of processing user rating data and generating relevant book recommendations based on the preferences of similar users. The implementation of Collaborative Filtering and K-Nearest Neighbor in a web-based sales system can provide more personalized recommendations and improve the user experience in product discovery.

Keywords— Machine Learning, Recommendation System, Collaborative Filtering, K- Nearest Neighbor

I. PENDAHULUAN

Pemanfaatan sistem rekomendasi memungkinkan pengguna memperoleh pilihan buku yang lebih sesuai melalui pengolahan riwayat penilaian yang tersedia. Sistem melakukan analisis terhadap pola preferensi pengguna untuk menghasilkan rekomendasi yang bersifat lebih spesifik. Dalam penelitian ini, pendekatan *Collaborative Filtering* yang dikombinasikan dengan algoritma *K-Nearest Neighbor* diterapkan untuk mengidentifikasi kesamaan karakteristik antar pengguna dalam menentukan buku yang berpotensi diminati [1].

Sebagai salah satu pendekatan pada sistem rekomendasi, *Collaborative Filtering* bekerja dengan memanfaatkan informasi rating pengguna terhadap suatu item. Data tersebut digunakan untuk mengukur kedekatan preferensi antar pengguna, sehingga buku yang disukai oleh pengguna dengan pola minat serupa dapat dijadikan rekomendasi bagi pengguna lainnya [2]. Meskipun demikian, tingkat ketepatan rekomendasi yang dihasilkan tetap bergantung pada jumlah serta kualitas data rating yang tersedia dalam proses analisis[3].

Salah satu algoritma yang banyak digunakan dalam sistem rekomendasi adalah *K-Nearest Neighbor*. Algoritma ini bekerja dengan mengklasifikasikan data berdasarkan kedekatan atau kemiripan karakteristik antar data. Prinsip kerja *K-Nearest Neighbor* adalah mencari sejumlah tetangga terdekat yang memiliki karakteristik serupa untuk menghasilkan rekomendasi yang sesuai dengan kebutuhan pengguna [4].

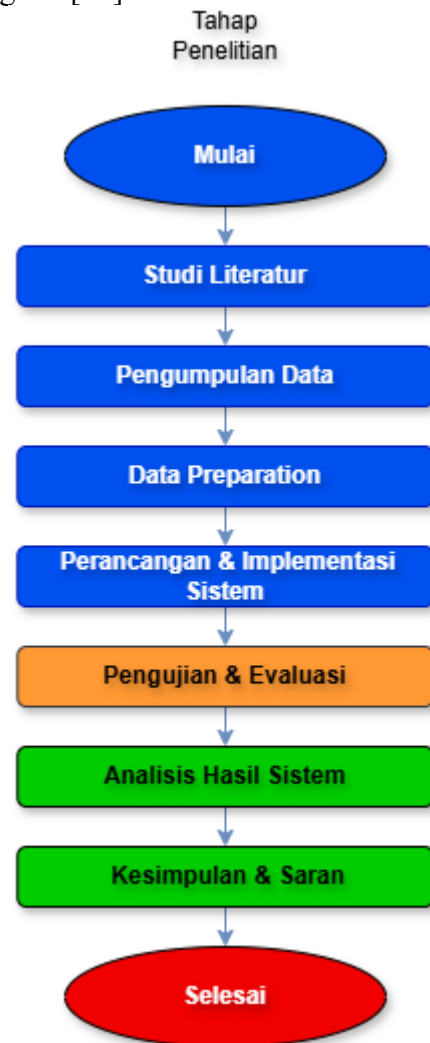
II. METODOLOGI PENELITIAN

Penelitian ini menggunakan metode *Research and Development* (R&D) yang bertujuan mengembangkan sistem rekomendasi buku berbasis web dengan menerapkan metode *User-Based Collaborative Filtering* dan algoritma *K-Nearest Neighbor* (KNN)[5]. Metode R&D dipilih karena tidak hanya berorientasi pada proses analisis, tetapi juga menghasilkan produk berupa sistem yang dapat diimplementasikan serta dievaluasi. [6]

Penelitian diawali dengan studi literatur untuk memperoleh landasan teori mengenai sistem rekomendasi, *User-Based Collaborative Filtering*, *Cosine Similarity*, algoritma KNN, serta teknologi pendukung yang digunakan dalam pengembangan sistem. Selanjutnya dilakukan pengumpulan data menggunakan *Book Recommendation Dataset* yang diperoleh dari platform Kaggle sebagai sumber data penelitian[7]. Dataset tersebut kemudian dipersiapkan melalui proses *data preparation* yang meliputi *data cleaning*, penanganan *missing value*, penghapusan data duplikat, seleksi atribut, serta pembentukan *user-item matrix* sebagai dasar dalam proses perhitungan rekomendasi. [8]

Proses berikutnya adalah perancangan dan implementasi sistem berbasis web menggunakan Laravel sebagai *framework*, MySQL sebagai basis data, dan Python sebagai bahasa pemrograman untuk menjalankan proses komputasi algoritma[9]. Sistem rekomendasi dibangun dengan menerapkan metode *User-Based Collaborative Filtering* yang memanfaatkan *Cosine Similarity* untuk menghitung tingkat kemiripan antar pengguna, sedangkan algoritma KNN digunakan untuk menentukan pengguna dengan tingkat kemiripan tertinggi sebagai dasar dalam memprediksi *rating* dan menghasilkan rekomendasi buku. Setelah sistem selesai dikembangkan, dilakukan pengujian

menggunakan metrik RMSE, MAE, Precision, Recall, F1-Score, Precision@10, dan Recall@10 untuk mengukur performa rekomendasi yang dihasilkan[10]. Tahap akhir penelitian dilakukan melalui analisis terhadap hasil rekomendasi serta verifikasi menggunakan perhitungan manual untuk memastikan bahwa implementasi algoritma telah sesuai dengan proses perhitungan yang dirancang, sehingga dapat diperoleh kesimpulan mengenai kinerja sistem rekomendasi yang dikembangkan.[11]



Gambar 1 Alur Penelitian

III. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan sistem rekomendasi buku berbasis web yang dikembangkan menggunakan metode *Research and Development* (R&D). Hasil penelitian disajikan berdasarkan tahapan penelitian yang telah dilakukan, dimulai dari pengumpulan dataset, *data preparation*, perancangan dan implementasi sistem, pengujian serta evaluasi, hingga analisis terhadap hasil rekomendasi yang dihasilkan. Penyajian hasil pada setiap tahapan bertujuan untuk menunjukkan proses pengembangan sistem secara sistematis sekaligus mengevaluasi kinerja metode *User-Based Collaborative Filtering* dan algoritma *K-Nearest Neighbor* (KNN) dalam menghasilkan rekomendasi buku yang relevan.

1. Pengumpulan Dataset

Tahap pengumpulan dataset dilakukan menggunakan **Book Recommendation Dataset** yang diperoleh dari platform Kaggle sebagai sumber data penelitian. Dataset terdiri atas tiga tabel utama, yaitu **Books**, **Users**, dan **Ratings**. Tabel **Books** berisi informasi buku, seperti ISBN, judul, penulis, tahun terbit, dan penerbit. Tabel **Users** memuat data pengguna, sedangkan tabel **Ratings** berisi data penilaian pengguna terhadap buku yang digunakan sebagai dasar dalam pembentukan rekomendasi.

Berdasarkan hasil pengumpulan data, penelitian ini menggunakan **1.391 data buku**, **445 data pengguna**, dan **5.871 data rating**. Seluruh data kemudian disimpan ke dalam basis data MySQL dan digunakan sebagai sumber informasi untuk membangun sistem rekomendasi berbasis *User-Based Collaborative Filtering* dan algoritma *K-Nearest Neighbor*.

Tabel 1. Dataset Penelitian

Dataset	Jumlah Data
Books	1.391
Users	445
Ratings	5.871

2. Data Preparation

Tahap *data preparation* dilakukan untuk menghasilkan dataset yang siap digunakan pada proses rekomendasi. Proses ini meliputi *data cleaning*, penanganan *missing value*, penghapusan data duplikat, serta seleksi atribut yang relevan. Atribut yang digunakan dalam penelitian meliputi **User-ID**, **Book-ID**, dan **Book-Rating**, karena ketiga atribut tersebut menjadi dasar dalam membangun hubungan antara pengguna dan buku.

Setelah proses persiapan data selesai dilakukan, data *rating* diubah ke dalam bentuk *user-item matrix* yang merepresentasikan hubungan antara pengguna dan buku berdasarkan nilai *rating* yang diberikan. Matriks ini digunakan sebagai masukan pada proses perhitungan *Cosine Similarity* untuk mengukur tingkat kemiripan antar pengguna sebelum menghasilkan rekomendasi buku.

Tabel 2. User-Item Matrix

User-ID	Book-ID 99	Book-ID 315	Book-ID 326	Book-ID 490	Book-ID 1306
277157	8	9	7	0	0
277427	8	0	0	0	7
278026	0	0	0	6	7
278137	0	6	8	10	0

Nilai 0 menunjukkan bahwa pengguna belum memberikan *rating* terhadap buku tersebut.

Misalnya menghitung kemiripan antara User 277157 dan User 278137.

Vektor *rating*:

$$U_{277157} = (8, 9, 7, 0, 0)$$

$$U_{278137} = (0, 6, 8, 10, 0)$$

Vektor *rating*:

$$Similarity(u, v) = \frac{\sum_{i=1}^n (R_{u,i} \times R_{v,i})}{\sqrt{\sum_{i=1}^n R_{u,i}^2} \times \sqrt{\sum_{i=1}^n R_{v,i}^2}}$$

Perhitungan Dot Product:

$$\begin{aligned} & (8 \times 0) + (9 \times 6) + (7 \times 8) + (0 \times 10) \\ & \quad + (0 \times 0) \\ & = 0 + 54 + 56 + 0 + 0 \\ & = 110 \end{aligned}$$

Panjang Vektor User 277157:

$$\begin{aligned} & \sqrt{8^2 + 9^2 + 7^2 + 0^2 + 0^2} \\ &= \sqrt{64 + 81 + 49} \\ &= \sqrt{194} \\ &= 13.93 \end{aligned}$$

Panjang Vektor User 278137:

$$\begin{aligned} & \sqrt{0^2 + 6^2 + 8^2 + 10^2 + 0^2} \\ &= \sqrt{36 + 64 + 100} \\ &= \sqrt{200} \\ &= 14.14 \end{aligned}$$

Nilai Cosine Similarity:

$$\begin{aligned} \text{Similarity} &= \frac{110}{13.93 \times 14.14} \\ &= \frac{110}{196.99} \\ &= 0.5584 \end{aligned}$$

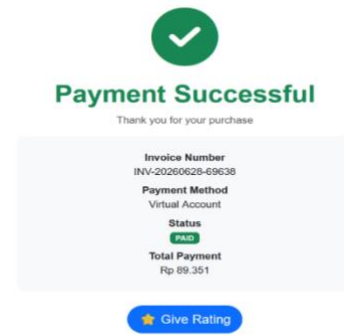
Hasil Akhir:

$$\text{Similarity}(277157, 278137) = 0.5584$$

3. Perancangan & Implementasi Sistem

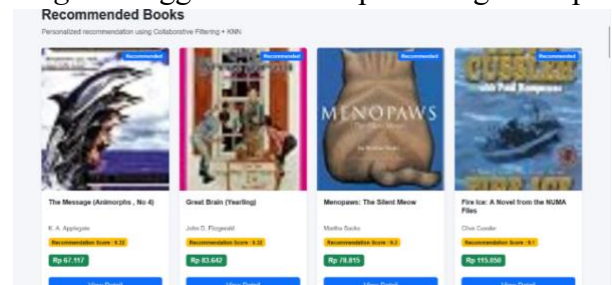
Tahap perancangan dan implementasi menghasilkan sistem rekomendasi buku berbasis web yang dikembangkan menggunakan Laravel sebagai *framework*, MySQL sebagai basis data, dan Python sebagai media komputasi algoritma. Sistem memanfaatkan data *rating* pengguna yang telah melalui proses *data preparation* untuk membentuk *user-item matrix* sebagai dasar perhitungan rekomendasi. Selanjutnya, metode *User-Based Collaborative Filtering* digunakan untuk menghitung tingkat kemiripan antar pengguna menggunakan *Cosine Similarity*, kemudian algoritma *K-Nearest Neighbor* (KNN) menentukan pengguna dengan nilai *similarity* tertinggi (*nearest neighbors*).

Hasil proses tersebut digunakan untuk memprediksi *rating* buku yang belum pernah dinilai oleh pengguna sehingga sistem mampu menghasilkan rekomendasi buku secara otomatis sesuai dengan preferensi pengguna.



Gambar 1. Halaman Pemberi Rating

Halaman pemberian *rating* digunakan pengguna untuk memberikan penilaian terhadap buku yang telah dibaca dengan skala 1–10. Data *rating* yang tersimpan pada basis data menjadi masukan utama dalam proses perhitungan *User-Based Collaborative Filtering* sehingga sistem dapat mengenali pola



Gambar 2. Halaman Rekomendasi Buku

preferensi setiap pengguna.

Setelah proses perhitungan selesai dilakukan, sistem menampilkan daftar Top-10 rekomendasi buku berdasarkan hasil prediksi *rating*. Rekomendasi yang ditampilkan merupakan buku dengan nilai prediksi tertinggi yang diperoleh dari proses *Collaborative Filtering* dan algoritma KNN, sehingga rekomendasi yang dihasilkan bersifat lebih personal dan relevan terhadap preferensi pengguna.

4. Pengujian dan Evaluasi

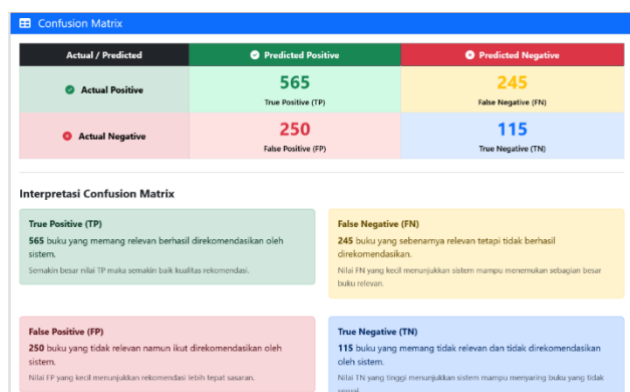
Tahap pengujian dilakukan untuk mengevaluasi performa sistem rekomendasi buku yang telah dikembangkan menggunakan metode *User-Based Collaborative Filtering* dan algoritma *K-Nearest Neighbor* (KNN). Pengujian dilakukan dengan mengukur tingkat kesalahan prediksi menggunakan *Root Mean Square Error* (RMSE) dan *Mean Absolute Error* (MAE), serta kualitas rekomendasi menggunakan metrik *Precision*, *Recall*, *F1-Score*,

Precision@10, dan *Recall@10*. Selain itu, analisis *Confusion Matrix* digunakan untuk mengetahui jumlah prediksi yang benar maupun salah dalam proses rekomendasi.



Gambar 3. Halaman Halaman Evaluasi Sistem

Berdasarkan hasil evaluasi pada gambar diatas sistem memperoleh nilai RMSE sebesar 1,64 dan MAE sebesar 1,36, yang menunjukkan bahwa tingkat kesalahan prediksi *rating* relatif rendah. Sementara itu, nilai *Precision* sebesar 69,33%, *Recall* sebesar 69,75%, dan *F1-Score* sebesar 69,54% menunjukkan bahwa sistem mampu menghasilkan rekomendasi dengan tingkat ketepatan dan keseimbangan yang cukup baik. Nilai *Precision@10* sebesar 76,70% mengindikasikan bahwa sebagian besar dari sepuluh rekomendasi teratas bersifat relevan, sedangkan *Recall@10* sebesar 98,90% menunjukkan bahwa hampir seluruh buku yang relevan berhasil direkomendasikan oleh sistem.



Gambar 4. Halaman Confusion Matrix

Berdasarkan hasil *Confusion Matrix*, sistem menghasilkan 565 data *True Positive* (TP), 115 data *True Negative* (TN), 250 data *False Positive* (FP), dan 245 data *False Negative* (FN). Dominasi nilai

True Positive dibandingkan *False Positive* menunjukkan bahwa sebagian besar rekomendasi yang diberikan telah sesuai dengan preferensi pengguna. Hasil ini juga mendukung nilai *Precision*, *Recall*, dan *F1-Score* yang diperoleh, sehingga implementasi metode *User-Based Collaborative Filtering* dan algoritma KNN mampu memberikan rekomendasi buku yang cukup akurat pada dataset penelitian.

5. Analisis Hasil Sistem

Analisis hasil sistem dilakukan untuk memastikan bahwa proses rekomendasi berjalan sesuai dengan tahapan yang telah dirancang. Gambar 6 menunjukkan daftar buku yang telah diberikan *rating* oleh pengguna aktif (*User-ID* 277157). Data *rating* tersebut menjadi masukan utama dalam pembentukan *user-item matrix* yang selanjutnya digunakan untuk menghitung tingkat kemiripan antar pengguna menggunakan metode *User-Based Collaborative Filtering* dengan *Cosine Similarity*.

Hasil perhitungan tersebut kemudian diproses menggunakan algoritma *K-Nearest Neighbor* (KNN) untuk menentukan pengguna dengan tingkat kemiripan tertinggi sebagai dasar dalam proses prediksi *rating* dan pembentukan rekomendasi buku.

```
=====
10 BUKU YANG DIRATING OLEH LOGIN USER
=====
```

book_id	title	rating
1263	As mulheres deviam vir com livro de instrução	1
253	Agent of Influence	1
454	The Virgin Suicides	
315	Chance	
327	Fire Ice: A Novel from the NUMA Files	
99	Transitions: Making Sense of Life's Changes	
1350	Taxi! : A novel	
830	Angels &	
1057	To the Hermitage	
746	Betrayal	

Gambar 5. Halaman Phyton Rekomendasi Buku

Berdasarkan data *rating* pada pengguna aktif telah memberikan penilaian terhadap beberapa buku dengan rentang nilai 8 hingga 10, yang menunjukkan tingkat ketertarikan pengguna terhadap buku-buku tersebut. Informasi tersebut dimanfaatkan sistem untuk mengenali pola preferensi pengguna dan membandingkannya dengan pola *rating* pengguna lain. Selanjutnya, dilakukan verifikasi melalui

perhitungan manual pada proses pembentukan *user-item matrix*, perhitungan *Cosine Similarity*, pemilihan *nearest neighbors*, serta prediksi *rating*. Sehingga implementasi metode *User-Based Collaborative Filtering* dan algoritma *K-Nearest Neighbor* telah berjalan sesuai dengan rancangan penelitian.

IV. KESIMPULAN

Penelitian ini menghasilkan sistem rekomendasi buku berbasis web dengan menerapkan metode *K-Nearest Neighbor* dan *User-Based Collaborative Filtering*. Sistem mampu menentukan rekomendasi berdasarkan pola kesamaan preferensi antar pengguna. Hasil pengujian menunjukkan bahwa metode yang digunakan memiliki tingkat akurasi yang baik dengan nilai kesalahan prediksi rendah, sehingga dapat membantu pengguna memperoleh pilihan buku yang sesuai dengan minatnya.

UCAPAN TERIMA KASIH

Penyelesaian penelitian ini tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, penulis menyampaikan apresiasi kepada Universitas Duta Bangsa Surakarta serta para dosen pembimbing yang telah memberikan bimbingan, arahan, dan saran selama pelaksanaan penelitian. Penulis juga mengucapkan terima kasih kepada pihak-pihak lain yang turut berkontribusi dalam mendukung hingga penelitian ini dapat terselesaikan.

REFERENSI

- [1] A. Waluyo dkk., "PENGEMBANGAN SISTEM BACKEND REKOMENDASI PRODUK DENGAN COLLABORATIVE FILTERING BERBASIS SINGULAR VALUE DECOMPOSITION," *Journal of Digital Business and Technology Innovation (DBESTI)*, vol. 2, no. 2, hlm. 182–190, 2025, [Daring]. Tersedia pada: <https://journal.nurulfikri.ac.id/index.php/DBESTI>
- [2] Ilyas Tonapa, Sudianto Lande, dan Hermin Arrang, "jurnal+ilyas," 2025.
- [3] A. S. Diantika, Y. Firmanto, A. J. Akuntansi, F. Ekonomi, dan D. Bisnis, "IMPLEMENTASI MACHINE LEARNING PADA APLIKASI PENJUALAN PRODUK DIGITAL (STUDI PADA GRABKIOS)," 2021.
- [4] Y. Ahda Wicaksono, F. Wahyu Christanto, dan I. Saufik Suasana, "Implementasi Algoritma K-Nearest Neighbor Berbasis Website pada Sistem Rekomendasi Persewaan Alat Camping," 2025. [Daring]. Tersedia pada: <https://jurnal.uniraya.ac.id/index.php/JI>

- [5] A. Rochmad Wahono, B. Aji Saputra, dan F. Fadlu Rahman, "Sistem Rekomendasi Film Menggunakan Metode Content-Based Filtering dan Algoritma K-Nearest Neighbors (KNN)," hlm. 18–2024, 2025.
- [6] C. Huda, Y. Heryadi, Lukas, dan W. Budiharto, "Smart Tourism Recommender System Modeling Based on Hybrid Technique and Content Boosted Collaborative Filtering," *IEEE Access*, vol. 12, hlm. 131794–131808, 2024, doi: 10.1109/ACCESS.2024.3450882.
- [7] D. Ezekiel Griffin, K. D. K. A. Wardani, dan A. A. N. E. S. Gorda, "Optimalisasi Inventaris UMKM di Denpasar melalui Penerapan Sistem Rekomendasi Pemesanan Berbasis Machine Learning," *Jurnal Pengabdian UntukMu NegeRI*, vol. 9, no. 3, hlm. 504–511, 2025, doi: 10.37859/jpumri.v9i3.9973.
- [8] L. Ph.D., F. Loekman, dan M. Augustine, "PENGEMBANGAN WEBSITE FUKU ELEKTRONIK SEBAGAI MEDIA PEMASARAN SECARA DIGITAL," *Jurnal Bakti Masyarakat Indonesia*, vol. 5, no. 3, 2023, doi: 10.24912/jbmi.v5i3.20444.
- [9] R. I. Salam dan D. E. Putra, "Perancangan dan Implementasi Website Toko Kue: Peningkatan Penjualan dan Kerlibatan Pelanggan," *Jurnal Pustaka Robot Sister (Jurnal Pusat Akses Kajian Robotika, Sistem Tertanam, dan Sistem Terdistribusi)*, vol. 2, no. 1, 2023, doi: 10.55382/jurnalpustakarobotsister.v2i1.448.
- [10] H. Hendra dan Y. F. Riti, "Perancangan Dan Implementasi Website Dengan Konsep Ui/Ux Untuk Mengoptimalkan Marketing Perusahaan," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 11, no. 3s1, 2023, doi: 10.23960/jitet.v11i3s1.3430.
- [11] Y. C. Hoar Siki, D. J. Manchat, E. Ngaga, dan Y. P. Bria, "Implementasi Website Penjualan Bagi Pelaku UMKM Desa Napan – Bikomi Utara," *Jurnal Pengabdian kepada Masyarakat Nusantara*, vol. 5, no. 1, 2024, doi: 10.55338/jpkmn.v5i1.2682.