Received May 1st, 2025; accepted June 29th, 2025. Date of publication June 30th, 2025
Digital Object Identifier: https://doi.org/10.25047/jtit.v12i1.5996

Machine Learning-Based Book and Library Recommendation Application Using Content-Based Filtering Method

NOVIT SAPUTRI[1], PULUT SURYATI[2]

[1,2] Sistem Informasi, Teknologi Informasi, Universitas Teknologi Digital Indonesia

CORESPONDING AUTHOR: NOVIT SAPUTRI (email : novitputria0113@gmail.com)

ABSTRACT This research focuses on the development of a Book Recommendation System application, which aims to simplify book searches based on reader preferences. Faced with the challenge of selecting books that match their interests, readers often find it difficult to navigate the vast array of information available. The availability of different genres and authors can make the selection process complicated and time-consuming. Therefore, the development of a Book Recommendation System application is a relevant and necessary solution. The author's research question is how to optimize the book search experience through accurate recommendation algorithms. This project will use a collaborative algorithm-based recommendation model to analyze reader behavior patterns and provide accurate book recommendations based on the similarity of reader preferences. The application that the author designed allows readers to easily find suitable books. The author is committed to overcoming readers' difficulties in dealing with the abundance of information, increasing the accessibility of books, and improving reader satisfaction. The success of the project is measured by the application's ability to provide satisfactory recommendations, find relevant works, and simplify the literature exploration process. With the development of the era of vast information, the development of a Book Recommendation System is a relevant solution to guarantee a more enjoyable and efficient reading experience for readers.

KEYWORDS : Application, Collaborative Algorithm, Reader Preferences, Recommendation System.

I. INTRODUCTION

In today's digital age, advances in information technology have had a significant impact on various aspects of life, including literacy and libraries. The increasing number of books available in physical and digital formats has created new challenges for readers, namely the difficulty of finding books that suit their interests and preferences. The wide range of genres, authors, and topics available makes the search process complex and often time-consuming. This calls for smart solutions that can help readers navigate information more efficiently and personally. One potential approach is the implementation of a technology-based recommendation system that can automatically analyze user preferences. With this system, the reading experience can be enhanced hrough more relevant recommendations, enabling readers to find books that meet their needs more quickly and easily.[1]

One approach that can be used to address this challenge is the development of a machine learning-based book recommendation system. This system is designed to analyze user preference data and automatically provide relevant book suggestions. Content-based filtering is one of the most effective techniques in recommendation systems because it can suggest new items based on the similarity of content to items previously liked by users. Most previous studies still have some limitations, such as focusing only on genre as a content attribute without considering descriptions or other textual information. In addition, there are also recommendation systems that only evaluate the accuracy of algorithms quantitatively without considering the user experience aspect in practice. Some systems also rely on a collaborative filtering

approach that depends on large amounts of user data, making them less effective if user data is limited or new users do not yet have a history of interaction.[2]

This research focuses on developing a book and library recommendation application that utilizes content-based filtering methods. This application aims to enhance the user experience in exploring literature, simplify the book search process, and strengthen the role of libraries as responsive information centers that meet readers' needs. By combining machine learning technology and book content data processing, this application is expected to provide more accurate and personalized recommendations for each user. The objective of this research is to present a relevant solution to the problem of book search, which is still commonly encountered in libraries and digital literacy services.[3]

In developing this application, data covering book information, user ratings, and content features such as genre, author, and book description were used. The data was analyzed using a content-based filtering algorithm that measures similarities between books based on these attributes. The results of this research are expected to contribute to the development of a more adaptive and user-centric smart technology-based library information system. This research differs from previous studies because it employs a more comprehensive content analysis, considers user experience aspects, and can be applied even with limited user data. With this approach, it is hoped that users can obtain book recommendations that better align with their interests in a more effective and efficient manner.

II. METHOD

Content-Based Filtering is a recommendation system method that suggests items (in this case books) to users based on content similarity between items. Recommendations are generated by analyzing the characteristics of an item that has been liked by the user, then searching for other items that have similar features to the item. This method does not require data from other users, as the main focus is on the features of the item itself.[4]

With this approach, book recommendation applications are able to provide relevant and personalized recommendations based solely on book content and individual user interactions, without the need for collective data from other users.[5]

In its implementation, the Content-Based Filtering method in this application uses text processing techniques to extract important features from book descriptions, such as genre, author, and frequently occurring keywords. Furthermore, vectorization algorithms such as TF-IDF (Term Frequency-Inverse Document Frequency) are applied to convert text information into numerical representations that can be compared with each other. The similarity matching process between books is performed using similarity measures, such as cosine similarity, to determine the extent to which two books have similar content. In this way, the system can automatically recommend books that have characteristics closest to the user's preferences.[6]

FIGURE 1. Content Based Filtering Method

These stages are designed to systematically address and achieve the research objectives and ensure robust and interpretable results.

1. Data Collection

The first step is to collect relevant data, including: Item data in the form of title, author, genre, description, year of publication, etc. User data in the form of ratings, previous interactions. This stage is important to ensure the availability of sufficient information as a basis for the recommendation process. [7]

2. Feature Extraction

Descriptive data from items is converted into numerical representations in order to be processed by Machine Learning algorithms. Text items such as title, author, or description are converted using the TF-IDF (Term Frequency-Inverse Document Frequency) method. The result is a feature vector for each item. This process allows the system to recognize patterns and relationships between items based on content similarity. [8]

3. Similarity Calculation

After each item is represented as a vector, the similarity between items is calculated using Cosine Similarity → measuring the angle between two vectors to determine how similar two items are. This step is central in determining the relevant books to recommend to the user.[9]

4. Recommendation Generation

Based on the similarity of items, the system will search for items that are most similar to those that the user has liked before. Sort and display recommendations based on the highest similarity value. In this way, users get a more personalized list

of recommendations that match their preferences.[10]

5. Evaluation Model

Evaluate how well the system provides recommendations that are relevant to subjective evaluation through user testing. This evaluation stage ensures that the system built is truly effective in improving the user experience.[11]

III. Results and Discussion

The result of this research is an integrated book recommendation application and library information system, which utilizes the Content-Based Filtering method as the main approach in providing book recommendations to users. This application is designed to make it easier for users, especially students or library members, to find books that match their interests based on data from books that have been borrowed or viewed previously. The system works by analyzing book metadata such as title, author, genre, and description, then calculating the level of similarity between books using cosine similarity. Thus, each user gets personalized and dynamic recommendation results.

Evaluation of the system was carried out through direct testing to users and analysis of the relevance of the recommendation results. Based on the results of limited testing, the system received positive feedback in terms of ease of use and accuracy of recommendations. The precision value shows that most of the recommended books are considered to be in accordance with the user's interests. However, the system still has some limitations, especially in terms of dependence on the completeness of book metadata. Therefore, in the future, the development of the system can be improved with a hybrid filtering approach that combines content-based and collaborative filtering methods to produce more accurate and comprehensive recommendations.

In addition, the integration of the library information system with this recommendation feature provides additional benefits for library managers in understanding user reading interest trends. Through interaction data and recommendation reports, the library can identify the most popular genres or types of books, which can be used as a basis for procuring new collections that are more targeted. This feature also helps increase book borrowing rates as users find it more helpful in finding relevant readings, thus encouraging a more active literacy culture in the school or library environment.

1. Data Collection

In this research, the data collection stage is an important part of building a content-based filtering-based book recommendation system. The data used includes attributes such as book title, author name, genre, short description, and year of publication. This information is needed to describe the characteristics of each book, so that the system can identify similarities between items based on these features. The main focus in this data collection is on the author attribute, so that recommendations can be provided by relying on author similarities between books. This approach was taken to facilitate the development of the initial model and simplify the process of calculating the similarity between items.[12]

The data sources used in this research are obtained from public datasets available on data sharing platforms such as Kaggle. The dataset used contains complete information about various book titles and their authors. This open source selection is done so that the data used is legal, freely accessible, and covers a wide range of genres and authors. By using this dataset, the development of a recommendation system can be done with representative data without having to collect data manually on a large scale. [13]

TABLE 1. Attribute Data Collection

Judul Buku	Penulis	Genre	Tahun Terbit
Harry Potter and the Socerer's Stone	J.K. Rowling	Fantasy	1997
The Hobbit	J.R.R. Tolkien	Fantasy	1937
To Kill a Mockingbird	Harper Lee	Fiction	1960
Pride and Prejudice	Jane Austen	Romance	1813

To support the transparency and validity of the research, the dataset used is attached as part of this document. The appendix contains basic information such as book title, author name, and several other relevant attributes. The dataset is used as the basis in the process of modeling a content-based filtering-based recommendation system that utilizes similarities between items. This appendix is also the main reference in the evaluation and testing of the developed recommendation system.[14]

2. Feature Extraction

Once the data has been collected, the next stage is feature extraction to convert the text data into a numerical representation so that it can be processed by machine learning algorithms. In this research, the method used is TF-IDF (Term Frequency-Inverse Document Frequency), which measures the importance of a word in one document compared to all documents in the dataset. Attributes such as the author's name are processed using TF-IDF to generate numerical vectors, where each vector represents the characteristics of each book based on relevant words. This process allows the system to calculate the degree of similarity between books using methods such as cosine similarity.[15]

The result of this feature extraction is TF-IDF vectors that are used to compare books with each other based on feature similarity. Using this representation, the recommendation system can

.

suggest other books that have similar content, not limited to author similarity, but also to the word patterns that appear.[16]

Term Frequency (TF)
Term frequency (TF) (1) can be defined as the relative frequency of a term (t) in a document (d).

The number is calculated by dividing the number of occurrences of the term in the document (f t , d) by the total number of terms in the document.(1)

$$tf\ (t,d) = \frac{f_{t,d}}{\sum_{t' \in f t' d}} \quad (1)$$

$$TF = \frac{Number\ of\ times\ the\ term\ appears\ in\ the\ document}{Total\ number\ of\ terms\ in\ the\ document} \quad (1)$$

Inverse Document Frequency (IDF)
Inverse document frequency (IDF) (2) measures the amount of information a term provides.

This value is calculated by dividing the total number of documents (N) by the number of documents containing the term. Then, taking the logarithm of the quotient.(2)

$$\text{idf}(t, D) = \log \frac{N}{|\{d \in D : t \in d\}|} \quad (2)$$

TF-IDF Formula
To calculate TF-IDF, (3) we need to multiply TF and IDF values (3)

$$\text{tfidf}(t,d,D) = \text{tf}(t,d) \cdot \text{idf}(t,D) \quad (3)$$

3. Similarity Calculation

After each item has been successfully represented as a vector through the TF-IDF method,(3) the next step is to calculate the degree of similarity between items. For this process, the cosine similarity method is used,(4) which is a measurement technique that calculates the angle between two vectors in multidimensional space. Cosine similarity does not depend on the length of the vectors, but rather focuses on the direction, so it is suitable for comparing text representations such as author names or book descriptions. The resulting similarity value ranges from 0 to 1, where a value of 1 indicates two perfectly identical items, and a value close to 0 indicates dissimilarity(4). [17]

$$\cos(\Theta ij) = \frac{\sum_k (d_{ik}\, d_{jk})}{\sqrt{\sum_k d_{ik}^2}\sqrt{\sum_k d_{jk}^2}} \quad (4)$$

Through cosine similarity calculation, the system can find books that have high closeness based on the extracted features. The results of this calculation are used to generate a recommendation list, where the books that have the highest similarity score will be recommended to the user. This process is the core of the content-based filtering-based recommendation system, as it enables the search for relevant books without the need for direct user interaction data.

TABLE 2. Cosine Similarity Calculation

Cosine Similarity Shape : (10000, 10000)

book_title	CCNA Exam Certification Guide (CCNA Exam 640-407)	Little Women (Scholastic Classics)	he Darwin Awards III: Survival of the Fittest	Comets (Issac Asimov's Wonderful World of Science Fiction, No 4)
book_title				
Hyperion	0.0	0.0	0.0	0.0
Babyhood	0.0	0.0	0.0	0.0
The BRIDE - PROMOTIONAL	0.0	0.0	0.0	0.0
The Hope: A Novel	0.0	0.0	0.0	0.0
C Is for Corpse (Kinsey Millhone Mysteries (Paperback))	0.0	0.0	0.0	0.0
Me and My Little Brain	0.0	0.0	0.0	0.0
My War	0.0	0.0	0.0	0.0
God on a Harley: A Spiritual Fable	0.0	0.0	0.0	0.0

With this approach, the system is able to provide accurate recommendations even if the user is using the application for the first time or has not done much interaction, which is known as the cold-start problem in recommendation systems. Since the main focus is on content similarity between books, the recommendations are stable and consistent as long as the complete book metadata is available. In

addition, this method also allows for further development, such as giving more weight to certain features (e.g. genre or description) to make the recommendations more aligned with the user's specific preferences.

4. Recommendation Generation

At the recommendation generation stage, the system uses a Content-Based Filtering approach to search for items that are most similar to the items that the user has previously liked. Data from each book is processed using the TF-IDF technique to convert information into numerical form, so that the degree of similarity can be calculated. Then, the system calculates the similarity between items using the cosine similarity method. Based on the similarity score, the system sorts and displays the top few items that have the highest similarity score with the user's selected item.

This sorting process ensures that the recommendations displayed are truly relevant and in accordance with the user's pre-recorded preferences. In addition, the system is also designed to display a variety of recommendations from various genres that still have a high degree of similarity, to prevent saturation due to recommendations that are too homogeneous. Thus, users not only get similar books, but also have the opportunity to explore new readings that are still within the scope of their interests. This approach aims to increase user satisfaction while encouraging more active engagement in reading activities.[18]

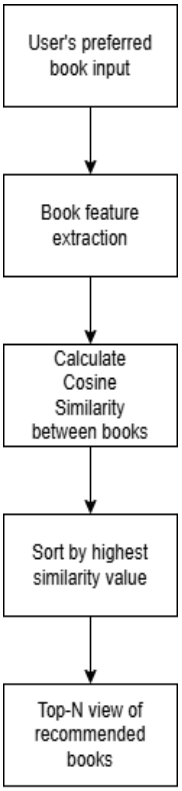

FIGURE 2. Flow of Recommendation Generation

The recommendation process in this system begins with user input in the form of books they like (user's preferred book input), which will be the main reference in searching for similar books. Next, the system performs feature extraction (book feature extraction) on the entire book collection by taking attributes such as title, author, genre, and description. These features are processed using text processing techniques and the TF-IDF (Term Frequency–Inverse Document Frequency) algorithm to convert textual information into numerical vector representations. Once all books are represented as vectors, the system calculates the similarity between books using the cosine similarity method, measured by the formula Cosine Similarity

$$(A, B) = \frac{\text{A.B}}{||\text{A}||x||B||}. \qquad (5)$$

This similarity value indicates how similar the content of the input book is to other books, where a value close to 1 means very similar. The system then sorts all books based on the highest similarity value (sort by highest similarity value) and displays the top-N view of recommended books that best match the user's preferences. This process is visualized in a modeling diagram illustrating the flow from user input, feature extraction, similarity calculation between vectors, to the final recommendation results based on content-based filtering.

The results of applying this method show that the system is able to provide recommendations that are relevant and in line with user preferences. The recommended items have similar characteristics to the initial item, thus increasing the likelihood of users being interested in the recommendation. Although it has produced fairly accurate recommendations, further development can be done by considering the addition of more features to enrich the recommendation results and expand the variety of items offered to users.

Algorithm 1 Modeling Content-Based Filtering

```
COMMAND (XT)
MODELING
books    = books[:10000]
ratings  = ratings[:5000]

tfidf    = TfidfVectorizer()
tfidf.fit(books.book_author)
# tfidf.get_feature_names_out()

RESULT (X)
TfidfVectorizer
TfidfVectorizer()
```

.

COMMAND **(XT)**
tfidf_matrix=
tfidf.fit_transform(books.book_author)
tfidf_matrix.shape

RESULT **(X)**
(10000, 5575)

COMMAND **(XT)**
tfidf_matrix.todense()

RESULT **(X)**

matrix([[0., 0., .., 0., 0.],
[0., 0., .., 0., 0.],
[0., 0., .., 0., 0.],
..,

Figure 5. *Modeling Content-Based Filtering*

5. Model Evaluation

Based on the test results using the evaluation matrix, the content-based filtering-based recommendation system shows a fairly good performance. The average Precision@10 value of 70% indicates that most of the books recommended by the system are indeed relevant to the user's preferences. In addition, the Recall@10 value of 58% shows that the system is able to find more than half of the relevant books available to the user. The F1-Score@10 value of 63% shows a balance between the accuracy and completeness of the recommendations generated, with the system being able to maintain good recommendation quality without sacrificing relevance coverage. [19]

Algorithm 2 Precision, Recall, dan F1-Score

COMMAND **(XT)**
# Tampilan hasil
import pandas as pd

hasil_df = pd.DataFrame(hasil_evaluasi).T
print (hasil_df)

RESULT **(X)**

	Precision@10	Recall@10	F1-Score@10
User 001	0.5	0.83	0.62
User 002	0.4	1.00	0.57
User 003	0.6	1.00	0.75

The following diagram helps the reader quickly see whether the model is better at precision or at capturing all positive cases (recall), as well as whether the two are balanced (F1-score).

Algorithm 3 Displaying Bar Chart

COMMAND **(XT)**
fig, ax = plt.subplots(figsize=(10,6))

rects1 = ax.bar(x - width, precision, width, label='Precision@10', color='skyblue')

rects2 = ax.bar(x, recall, width, label='Recall@10', color='lightgreen')

rects3 = ax.bar(x + width, f1_score, width, label='F1-Score@10', color='salmon')

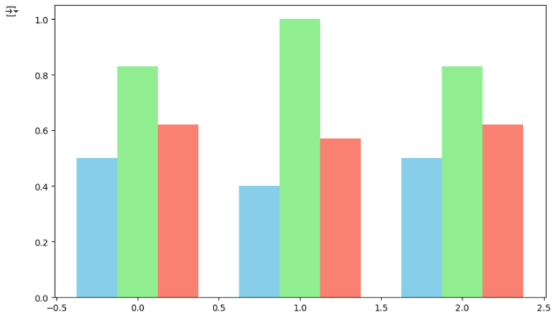

FIGURE 3. Bar Chart

Although the results of this quantitative evaluation show positive performance, there is room for improvement, especially in increasing recall so that more relevant books can be reached by the system. This shows that the model is effective in understanding the characteristics of books that users like, but it can still be further developed, for example by enriching content features or incorporating other approaches. This evaluation is an important basis that the system is working in the direction of development, but optimization is still needed to achieve maximum recommendations.

Algorithm 4 Recommendation Results

RESULT **(X)**
->> Rekomendasi 10 Buku Berdasarkan Penulis (Author) A-Z:

Book Title: On Her Own Ground : The Life and Times of Madam C.J. Walker
Author: A'Lelia Bundles
ImageURL:
http://images.amazon.com/images/P/0684825821.01.LZZZZZZZ.jpg

Book Title: The Dragon and the Unicorn
Author: A. A. Attanasio
ImageURL:
http://images.amazon.com/images/P/0061057797.01.LZZZZZZZ.jpg

Book Title: Winnie Ille Pu: A Latin Version of A. A. Milne's Winnie-The-Pooh
Author: A. A. Milne
ImageURL:
http://images.amazon.com/images/P/0525483357.01.LZZZZZZZ.jpg

Book Title: Pooh (Giant Board Book)
Author: A. A. Milne
ImageURL:
http://images.amazon.com/images/P/0525462325.01.LZZZZZZZ.jpg

The mobile design display is made to make it easier for readers to understand the application made by the author. Some mobile designs display important pages (sign in, home, search, book details, load confirmation, reservation success, and user).

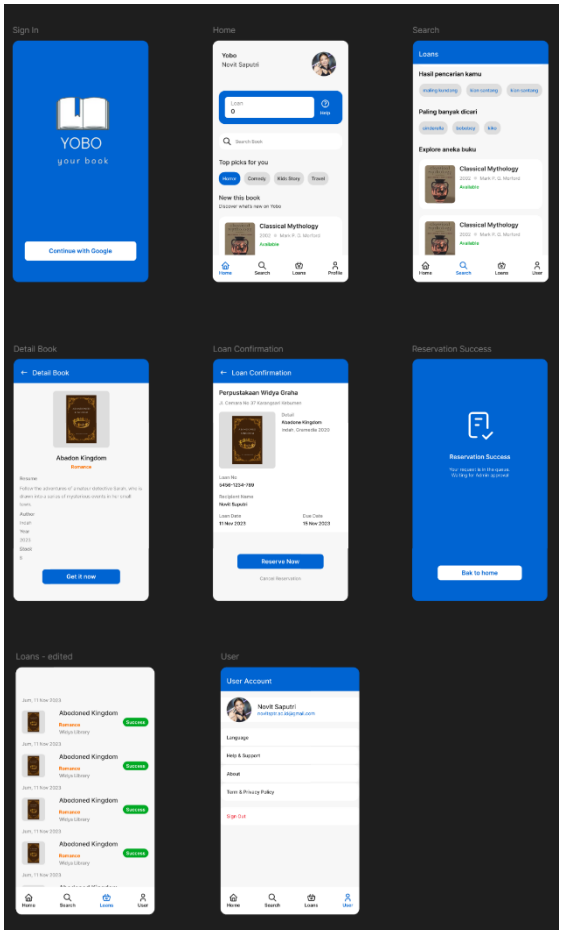

FIGURE 4. Mobile App Design

IV. CONCLUSION

The content-based filtering book recommendation application developed in this study has proven capable of providing relevant and personalized recommendations to users. By analyzing book content features such as genre, author, and description using TF-IDF and cosine similarity algorithms, the system can present book selections that match each user's preferences without requiring historical data from other users. The implementation of this system demonstrates improved efficiency in book search, particularly in digital library environments, and supports the transformation of libraries into more adaptive and responsive literacy centers tailored to individual information needs.

ACKNOWLEDGMENT

The author would like to express his deepest gratitude to PT.Dicoding Akademi Indonesia for their generous financial support, which was crucial in enabling the successful completion of the author's research on Machine Learning Based Book and Library Recommendation Application Using Content-Based Filtering Method. This funding has provided the necessary resources to advance the author's research and explore innovative solutions to the challenges faced by the technology industry and digital education in Indonesia. The unwavering trust and support from PT.Dicoding Akademi Indonesia has not only facilitated the development of this research, but also fostered a spirit of innovation and academic excellence. We greatly appreciate their commitment

to empowering researchers and driving impactful progress in this field.

REFERENCE

[1] S. Rakesh, "Movie Recommendation System Using Content Based Filtering," *Al-Bahir J. Eng. Pure Sci.*, vol. 4, no. 1, 2023, doi: 10.55810/2313-0083.1043.

[2] T. Ridwansyah, B. Subartini, and S. Sylviani, "Penerapan Metode Content-Based Filtering pada Sistem Rekomendasi," *Math. Sci. Appl. J.*, vol. 4, no. 2, pp. 70–77, 2024.

[3] A. Rianti, N. W. A. Majid, and A. Fauzi, "Machine Learning Journal Article Recommendation System Using Content Based Filtering," *JUTI J. Ilm. Teknol. Inf.*, vol. 22, no. 1, pp. 1–10, 2024.

[4] F. Maulana and E. B. Setiawan, "Performance of Deep Feed-Forward Neural Network Algorithm Based on Content-Based Filtering Approach," *INTENSIF J. Ilm. Penelit. dan Penerapan Teknol. Sist. Inf.*, vol. 8, no. 2, pp. 278–294, 2024, doi: 10.29407/intensif.v8i2.22904.

[5] H. Haryanto and A. Harits, "Otomatisasi Konten Statis dan Perilaku Dinamis dari Web E-Commerce Menggunakan Metode Content Based Filtering PT Multi Sinar Teknik," *sudo J. Tek. Inform.*, vol. 3, no. 3, pp. 106–112, 2024, doi: 10.56211/sudo.v3i3.421.

[6] A. A. Huda, R. Fajarudin, and A. Hadinegoro, "Sistem Rekomendasi Content-based Filtering Menggunakan TF-IDF Vector Similarity Untuk Rekomendasi Artikel Berita," *Build. Informatics, Technol. Sci.*, vol. 4, no. 3, pp. 1679–1686, 2022, doi: 10.47065/bits.v4i3.2511.

[7] A. Ikhwan and Z. Khalilah, "Sistem Informasi Pengolahan Data Rekomendasi Teknis Berbasis Web," *sudo J. Tek. Inform.*, vol. 2, no. 1, pp. 1–10, 2023, doi: 10.56211/sudo.v2i1.192.

[8] Y. Sari, A. R. Baskara, P. B. Prakoso, N. Royani, U. L. Mangkurat, and P. Korespondensi, "Perbandingan Metode Pembobotan Tf-Rf Dan Tf-Idf Dengan Dikombinasikan Dengan Weighted Tree Similarity Comparison of Tf-Rf and . Weighting Methods Tf-Idf With Combination With Weighted Tree Similarity for a Book," vol. 9, no. 3, 2022, doi: 10.25126/jtiik.202295709.

[9] S. Purwaningrum, A. Susanto, and A. Kristiningsih, "Synonym Recognition Influence in Text Similarity Detection Using Winnowing and Cosine Similarity," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 12, no. 3, pp. 219–226, 2023.

[10] E. K. A. Rahmawati, G. S. Nurohim, C. Agustina, and Z. Muttaqin, "Development of Machine Learning Model to Predict Hotel Room Reservation Cancellations," vol. 11, no. 2, 2024.

[11] A. Trisnadoli and D. Yanti, "Analysis of Prototyping Methods Implementation for Library Information Systems Development in SMAN 3 Bengkalis," *J. Nas. Pendidik. Tek. Inform.*, vol. 11, no. 3, pp. 280–289, 2022, doi: 10.23887/janapati.v11i3.50160.

[12] K. Febriani and K. Febriani, "Prediksi permintaan batu bara menggunakan machine learning," pp. 65–77, 2022.

[13] D. R. P. Noordi, H. Hasanah, and S. Sumarlinda, "Marvel Movie Recommendation System Using Hybrid Item-Based and Content-Based Filtering Methods," *TIERS Inf. Technol. J.*, vol. 5, no. 1, pp. 13–19, 2024, doi: 10.38043/tiers.v5i1.5209.

[14] A. Felfernig, M. Wundara, T. N. T. Tran, V. M. Le, S. Lubos, and S. Polat-Erdeniz, "Sports recommender systems: overview and research directions," *J. Intell. Inf. Syst.*, vol. 62, no. 4, pp. 1125–1164, 2024, doi: 10.1007/s10844-024-00857-w.

[15] N. P. Yunita, "Aplikasi Pencarian Hadis Menggunakan Vector Space Model Dengan Pembobotan TF-IDF Dan Confix-Stripping Stemmer," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 3, pp. 665–676, 2023, doi: 10.25126/jtiik.20231036736.

[16] S. A. Gunarto, E. S. Honggara, and D. D. Purwanto, "Website Sistem Rekomendasi dengan Content Based Filtering pada Produk Perawatan Kulit," *J. Sist. dan Teknol. Inf.*, vol. 11, no. 3, p. 399, 2023, doi: 10.26418/justin.v11i3.59049.

[17] V. N. November, "pemrograman Machine Learning , penulis memutuskan untuk melakukan studi Academy dan menyelesaikan proyek dengan judul ' MOVIEMU ' . Di era digital," vol. 4, no. 2, pp. 22–32, 2024.

[18] C. H. Saputra, "Integrasi Audit dan Teknik *Clustering* untuk Segmentasi dan Kategorisasi Aktivitas Log," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 1, pp. 209–214, 2024, doi: 10.25126/jtiik.20241118071.

[19] N. W. Utami, I. K. R. Arthana, and I. G. M. Darmawiguna, "Evaluasi Usability Pada E-Learning Universitas Pendidikan Ganesha Dengan Metode Usability Testing," *J. Nas. Pendidik. Tek. Inform.*, vol. 9, no. 1, p. 107,

2020, doi: 10.23887/janapati.v9i1.23663.

NOVIT SAPUTRI is an active student of the Bachelor of Information Systems Program, Faculty of Information Technology, Digital Indonesia University of Technology. In 2024, he participated in the Merdeka Belajar Kampus Merdeka (MBKM) Batch 2 program and was selected as a participant in Bangkit Academy, an independent study program resulting from the collaboration of Kemendikbudristek with Google, GoTo, and Traveloka. In this program, he took the Machine Learning learning path organized by PT Dicoding Akademi Indonesia. During the program, he learned various machine learning concepts in theory and practice, and developed a final project (Capstone Project) which was then used as the basis for writing this journal.

PULUT SURYATI, Lecturer at the Faculty of Information Technology, University of Digital Indonesia (UTDI). Who is also Novit Saputri's supervisor in completing this final project research. Pulut Suryati has several articles related to Computer Science.