Design and Development of a Soccer Shoes Recommendation Application Using NLP Model Implementation and Content-Based Filtering

Fikri Ananda Gusti Rakha[1], Rizky Ardiansyah[2*], Muhammad Syirajuddin Suja'i[3]

[1,2] Digital Telecommunication Network Study Program, Department of Electrical Engineering, State Polytechnic of Malang, 65141, Indonesia.

[3]Telecommunication Engineering Study Program, Department of Electrical Engineering, State Polytechnic of Malang, 65141, Indonesia

[1]rakhafikryy@gmail.com, [2]rizkyardiansyah@polinema.ac.id, [3]syirajuddin@polinema.ac.id

***Abstract*— Public interest in sports is increasing, including football, which is also supported by the development of sports facilities. In football, shoes are an important piece of equipment that must meet high quality standards and provide comfort when worn. Comfortable shoes can influence a player's performance and reduce the risk of injury. To prevent risks resulting from choosing inappropriate shoes, a solution is needed by building a soccer shoe recommendation application. This research aims to develop a web-based application that uses Natural Language Processing (NLP) technology. This technology functions to understand user preferences based on the answers given in the question and answer system and the search feature. This recommendation system combines the Content-Based Filtering method with NLP. The Content-Based Filtering method recommends shoes based on characteristics that users have liked before, while NLP is used to analyze product descriptions, user reviews and their needs. The research results show that this application is able to provide relevant shoe recommendations with 87% accuracy. In addition, testing proves that the integration of Natural Language Processing and Content-Based Filtering makes the process of getting recommendations and choosing shoes more efficient and increases user satisfaction. This application is expected to help soccer players choose shoes that suit their needs.**

Keywords*— *Content-Based Filtering, NLP, Preferences, Soccer Shoes, Users.

I. INTRODUCTION

In the current era of globalization, business competition is increasingly intense, especially with the rapid advancement of technology. Companies must be able to produce quality products to stay competitive, one of the areas with tight business competition is sports. In the modern era, sports have become an important need for every individual. Exercising not only maintains body health and reduces the risk of various diseases that can hinder activities, but also becomes a way for people to spend their free time, especially for those who have a hobby of exercising. The increase in public interest and awareness of sports today is followed by the development of sports facilities, especially in football. Talking about football, it feels like it is not far from football shoes, football shoes themselves must meet high quality standards with a design that can provide comfort and optimal performance on the field, choosing shoes that are comfortable on the feet is also very important for a player when they play during a match because it will affect performance and reduce the risk of player injury [1]. The model of the shoe sole plays an important role in the performance of players during training or competing on the field, the type of field used also has various types such as fields with grass, hard soil, and soft soil conditions [2].

The main problem faced by many buyers of football shoes is the difficulty in choosing the right shoes according to their needs and preferences. Although there are various choices of shoes on the market, buyers often feel confused because they cannot judge for sure whether the shoes will be comfortable to wear or fit their playing position. Buyers may also have difficulty understanding the various features of shoes, such as material, sole type, or design, which can affect performance and comfort during the match. For this reason, innovative features are needed provided by online platforms, including a recommendation system feature that provides various descriptions of soccer shoes regarding materials, soles and shoe models, as well as shoe recommendations according to the player's position. This plays an important role in attracting consumer interest and increasing product sales in stores, many buyers feel dissatisfied with the shoes they have purchased because they do not match their wishes. With this, the author offers a solution by creating an application based on a soccer shoe recommendation system to match buyers with products that are relevant to their interests, needs, and characteristics. So as to reduce the feeling of regret for buyers when they have bought a soccer shoe product but it does not match their wishes.

The concept of a recommendation system uses the content-based filtering and natural language processing (NLP) methods. Which content-based filtering method suggests items to users based on the characteristics or attributes of the items that they have previously liked [3]. Meanwhile, NLP is used to analyze and extract important features from product descriptions, user reviews, and preferences, allowing for a deeper understanding of user needs. When combined, NLP can provide a more accurate understanding of user preferences, while content-

*Corresponding author

based filtering can use this information to offer highly personalized and relevant recommendations. Content-based filtering (CBF) was chosen because it leverages detailed information about items that users like, making recommendations relevant. In contrast, collaborative filtering (CF) was not chosen because it relies on the user's interaction history, which makes it difficult to provide recommendations if user data is lacking or insufficient[4]. This combination allows for a more holistic and accurate recommendation process, which ultimately improves user satisfaction and the overall performance of the recommendation system.

This study aims to develop a web-based application that implements Natural Language Processing (NLP) technology to understand user preferences based on answer input in a question and answer system and also a search system according to user desires. In addition, this recommendation application system also implements the Content-Based Filtering algorithm applied to provide recommendations for soccer shoes that match the characteristics and needs of the user. The focus of this study is to improve the accuracy and relevance of shoe recommendations based on user preferences, making it easier for them to choose the shoes they want.

II. METHOD

A. Type of Research

This study is a development of previous studies, focusing on the implementation of shoe recommendation features in a soccer shoe recommendation application. The methods used include natural language processing (NLP) models, automatic image output technology, and content-based filtering techniques. This approach allows statistical analysis of system performance and measures user satisfaction based on their interactions with the application. The implementation of content-based filtering aims to provide more personalized and relevant recommendations according to the characteristics of the products that users are interested in.

B. System Design

In Fig. 1 there is a system block diagram as follows:

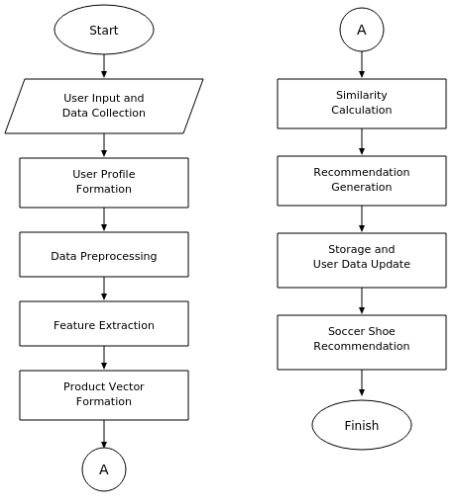

Figure 1. System Block Diagram

Fig. 2 illustrates the Natural Language Processing (NLP) block diagram that describes the sequential stages involved in processing textual data. The process begins with the Data Input Layer, where raw text data are collected from user input or document sources. The text then enters the Processing Layer, in which initial preprocessing steps such as tokenization, normalization, stop-word removal, and stemming are performed to clean and standardize the data. Next, the Feature Extraction Layer transforms the processed text into meaningful linguistic features by identifying important terms or patterns. These features are then converted into numerical representations in the Vector Formation Layer, allowing textual data to be processed mathematically. The resulting vectors are combined with existing user or system information in the Profile Integration Layer to enrich contextual understanding. Finally, the Evaluation and Feedback Layer assesses the performance of the NLP process and incorporates user feedback or system results to improve accuracy and effectiveness before the process concludes at the Finish stage.

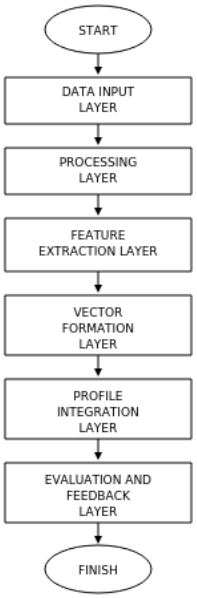

Figure 2. Natural Languange Procesing Block Diagram

In Fig. 3 shows the content-based filtering process in the form of a block diagram that illustrates the sequential workflow of the recommendation system. The process begins with system initialization, followed by the User Interaction Layer, where users provide input such as preferences, ratings, or item selections that reflect their interests. These interactions are captured in the Data Collection Layer, which gathers and stores relevant user behavior data and item attributes for further processing. The collected information is then processed in the Profile Building Layer to construct a user profile that represents individual preferences based on selected features or weighted attributes. Next, the Similarity Calculation Layer compares the user profile with available items in the system using similarity measurement techniques to identify items that closely match the user's interests. The results of this comparison are forwarded to the Recommendation Layer, where items are ranked and selected according to their

similarity scores. Finally, the recommended items are displayed to the user through the User Interface Layer, completing the recommendation process and enabling continuous personalization based on subsequent user interactions.

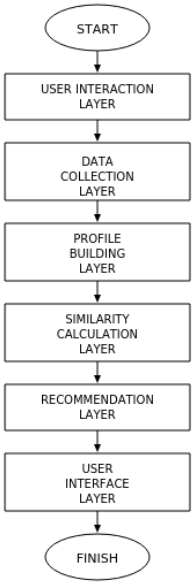

Figure 3. Content-Based Filtering Process Block Diagram

Fig 4 illustrates the Planning System block diagram that describes the workflow of the soccer shoe recommendation feature within the application. The process begins with system initialization, followed by user access to the soccer shoe recommendation feature. The system then evaluates whether the user selects a Q&A-based recommendation approach. If the user chooses this option, the system directs them to a chatbot-based recommendation system, where recommendations are generated based on the user's preferences and responses. Alternatively, if the user selects the searching recommendation system, the application displays recommendation results based on user-input keywords. In both paths, the application presents suitable soccer shoe recommendations tailored to the selected method. The system then evaluates whether the displayed recommendations align with the user's preferences. Once the user is satisfied, they select a recommended soccer shoe provided by the system and proceed to complete the payment transaction. The process concludes after the transaction is successfully completed, marking the end of the planning and recommendation workflow.

The flowchart in Fig. 5 illustrates the overall system operation starting from user authentication to the selection of the recommendation feature. The process begins with system initialization, after which the system checks whether the user already has an account. If the user has an existing account, they proceed to the login process by entering their username and password. Upon successful authentication, the user is directed to the recommendation feature selection menu. If the user does not have an account, the system directs them to the registration process, where they are required to enter a name and password. The system then processes the registration data and verifies whether the registration is successful. Once registration is completed successfully, the user can proceed to access the recommendation feature selection menu.

After authentication, the user selects one of two available recommendation methods, represented by connectors A and B. Connector A leads to the Q&A-based recommendation feature, where users answer a series of questions related to their preferences. The system analyzes the user's responses and processes them to generate suitable soccer shoe recommendations, which are then displayed to the user. Connector B leads to the search-based recommendation feature, where users input specific keywords. Based on these keywords, the application displays a list of recommended shoes that match the search criteria. In both approaches, users review the recommendations and select their preferred soccer shoes, after which the process concludes.

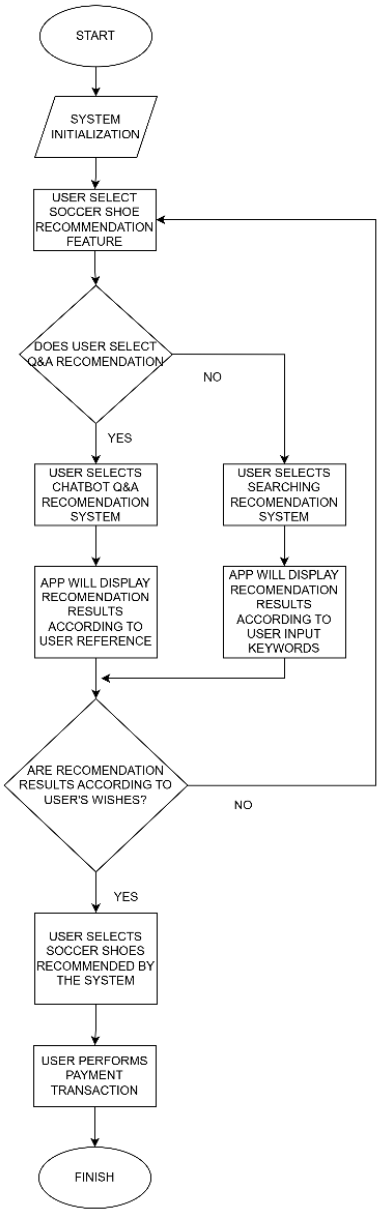

Figure 4. Planning System Block Diagram

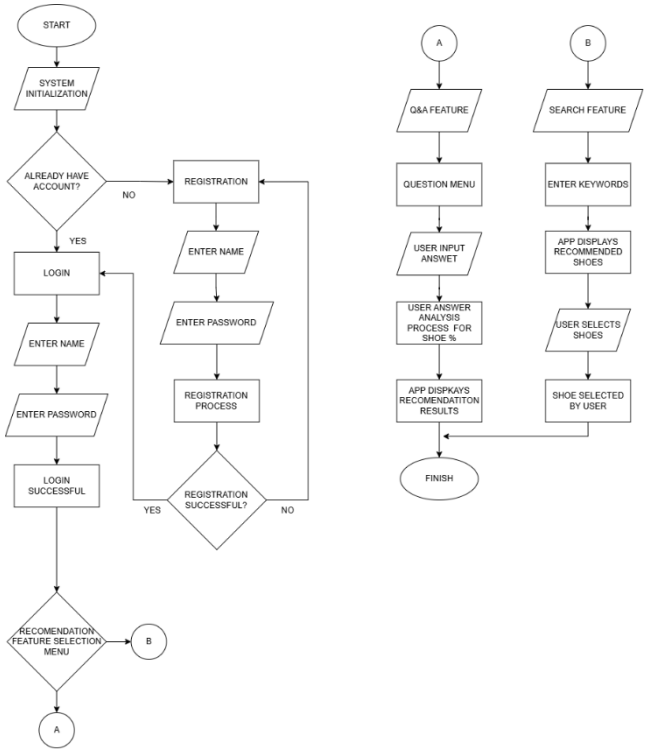

Figure 5. Application Planning Block Diagram

III. RESULTS AND DISCUSSION

A. System Functional Testing

In this research, functional testing of the system was carried out by the author by running the application in detail across each menu in the soccer shoe recommendation application system to determine which features functioned correctly and which did not. This testing aimed to ensure that all system functionalities operated as intended and that the application performed reliably. Functional testing was conducted on all main pages, including the login page, dashboard, key feature pages, and transaction page. Each feature was evaluated by comparing the expected results with the actual outcomes obtained during testing. The results of the functional testing are presented in tabular form to provide a clear and structured overview of system performance. Based on the testing results, all tested features showed outcomes that matched the expected behavior, indicating that the application runs properly and is ready for user use.

TABLE I
FUNCTIONAL TESTING OF THE APPLICATION LOGIN PAGE

No	Menu	Expected results	Results Obtained	Status
1.	Sign Up Form	Users can create a new account with valid data	Users can create a new account with valid data	Succeed
2.	Sign In Form	Users can enter the application with an existing account	Users can enter the application with an existing account	Succeed

TABLE II
APPLICATION DASHBOARD PAGE TESTING TABLE

No	Menu	Expected results	Results Obtained	Status
1.	Application Dashboard	The system displays main information according to user data	Successfully displays information according to user data	Succeed
2.	Menu Navigation	Users can enter the application with an existing account	Users can enter the application with an existing account	Succeed

TABLE III
KEY FEATURE PAGE FUNCTIONAL TESTING

No	Menu	Expected results	Results Obtained	Status
1.	Chatbot Feature Form	Thequestion and answer feature works well and the system provides recommend-dations according to user preferences	Successfully provided 15 product recommend-dations with the highest similarity	Succeed
2.	Search Feature Form	The search feature works well and the system displays products that match the search keywords	Successfully display relevant products based on keywords	Succeed

TABLE IV
TRANSACTION PAGE FUNCTIONAL TESTING

No	Menu	Expected results	Results Obtained	Status
1.	Cart Form	The system adds products to the shopping cart	The system adds products to the Shopping cart	Succeed

2.	Checkout Form	The system processes transactions with valid details	The system processes transactions with valid details	Succeed
3.	Transaction History	The system displays a list of transactions that the user has carried out	The system displays a list of transactions that the user has carried out	Succeed

B. User Acceptance Test (UAT) Testing with Questionnaires

The User Acceptance Test (UAT) was carried out to evaluate the level of user acceptance of the soccer shoe recommendation system that has been developed. This test involved 20 users, consisting of several members of the general public who are hobbies and want to find soccer shoes. Table 5 is the test scenario used:

TABLE V
SYSTEM EVALUATION QUESTIONNAIRE

No	Category	Question
1	Systems	Is the application display easy to recognize?
2		Is the application in the recommendation system easy to operate?
3		Does the color layout in the app help understand the information easily?
4		Is the application response speed adequate for user needs?
5		Is product-related information displayed in a logical order and easy to understand?
6	User	Is the menu display in this recommendation application easy for users to understand?
7		Are the pages that users need easy to find?
8		Are the text and information displayed easy to read?
9		Are the application features in the form of chatbot features and search features easy to access?
10		Are the chatbot features and search features able to provide recommendations according to the user's wishes?
11		Is the menu display in this recommendation application easy for users to understand?
12	Interaction	Do interaction features, such as buttons or menus, work well when accessing product information?
13		Do in-app product descriptions match user expectations?
14		Do users find the experience of using this application enjoyable?
15		Does the application provide adequate confirmation when users complete certain processes, such as transactions or product selection?

Table 6 is a table of questionnaire results answered by 20 respondents based on the test and conclusion tables for each category:

TABLE VI
QUESTIONNAIRE RESULTS

Category	Total Score	Ideal Score	Percentage (%)	Conclusion
Category 1: Systems	405	500	81%	Based on the evaluation, the application meets the ease of use aspect with an achievement of 81%. The layout is clear, responses are fast, and product information is structured. Although adequate, development is still needed for a more optimal experience.
Category 2: Users	486	600	81%	Based on the evaluation, the application meets ease of use with an achievement of 81%. The layout is clear, the features are easy to access, and the response is adequate. Product information is structured

Category	Total Score	Ideal Score	Percentage (%)	Conclusion
				so that it is easy to understand, but can still be developed further
Category 3: Interaction	334	400	83.5%	Based on the evaluation, the application meets the interaction aspect with an achievement of 83.5%. Interactions run well, features function optimally, and responses are as needed. However, Development is still needed to improve the quality of interaction.

C. Questionnaire Test Result Diagram for Each Category

The system category questionnaire results reveal that 51% of respondents rated the system as Neutral (C), 33% as Agree (S), 15% as Disagree (TS), and 10% as Strongly Agree (SS), as shown in Fig. 6.

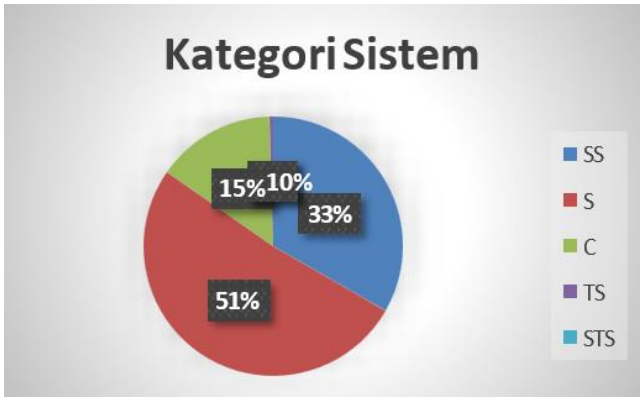

Figure 6. System Category Questionnaire Testing Diagram

The user category questionnaire results indicate that 47% of respondents rated the system as Agree (S), 36% as Strongly Agree (SS), 17% as Neutral (C), and 0% as Disagree (TS), as shown in Fig. 7.

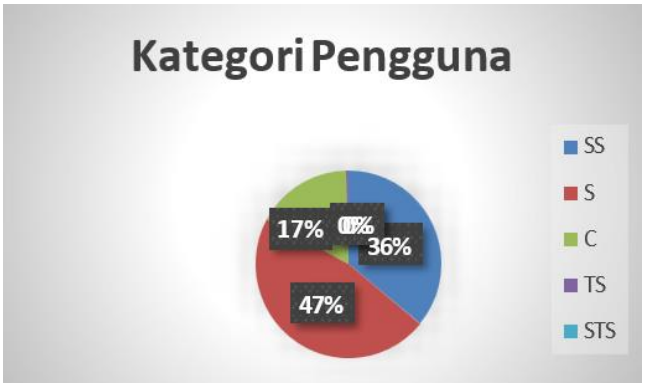

Figure 7. User Category Questionnaire Testing Diagram

The interaction category questionnaire results show that 45% of respondents rated the system as Strongly Agree (SS), 41% as Agree (S), 14% as Neutral (C), and 0% as Disagree (TS), as shown in Fig. 8.

Figure 8. Interaction Category Questionnaire Testing Diagram

D. Analysis of Questionnaire Testing Result for Each Category

In the System Category, the dominance of Agree (S) answers of 84.5% indicates that the system has met most of the respondents' expectations, especially in terms of functionality and reliability. This value indicates the system's superiority in providing stable performance and according to user needs. However, 15% of respondents who chose Sufficient (C) indicated that there were several aspects that needed to be improved, such as speed or additional features that were considered important by a small number of users. Meanwhile, the very small percentage of Disagree (TS) answers of 0.5% indicated that the majority of respondents did not experience major problems with the system.

In the User Category, the majority of respondents chose Agree (S) and Strongly Agree (SS) with a total of 83%, indicating that the user's ease and comfort aspects were considered very good. This value reflects the system's advantages in terms of an intuitive interface, easy-to-understand navigation, and a positive user experience. However, 17% of respondents chose Sufficient (C), which is a

sign that several parts, such as interface design or ease of access to certain features, need to be improved to further improve user comfort.

In the Interaction Category, the dominance of Strongly Agree (SS) answers of 65% indicates that the interaction between users and the system is running very well, especially in terms of responsiveness and process efficiency. This value shows the system's advantages in providing a fast and precise interaction experience. However, 14.4% of respondents chose Sufficient (C), which means there is still room for improvement, especially in increasing accessibility or optimizing the system to make it easier for all users .

E. Hosting Testing

The hosting testing results demonstrate that the system is capable of handling more than 25 shoe data entries with attributes and images from 5 sellers simultaneously, and successfully accommodates 45 users per day without any problems, as shown in Table 7.

TABLE VII
HOSTIME TESTING RESULT

No	Test Type	Results
1.	Maximum Data Testing	Capable of handling more than 25 shoe data along with attributes and images that can be entered by 5 sellers simultaneously
2.	Simultaneous User Testing	Testing was carried out with 45 users a day, all of whom successfully logged in and used the system without problems

F. Analysis of Hosting Test Result

The test results show that the server used is able to handle the amount of data and users simultaneously well. The hosting used in the form of Jagoan Hosting has been proven to be able to handle more than 25 shoe data along with attributes and images entered by 5 sellers simultaneously. Apart from that, testing with 45 users in one day showed that the system could operate without problems, with all users successfully logging in and using the system smoothly. Thus, it can be concluded that the hosting used is capable enough to support the needs of this application on the current scale and is ready to be implemented directly.

IV. CONCLUSION

This study successfully developed a web-based soccer shoe recommendation system by integrating Natural Language Processing (NLP) and Content-Based Filtering (CBF) methods. The NLP technique processes user input through tokenization, stop word removal, and stemming, while word2vec generates word vectors for analyzing product descriptions and user preferences. The CBF model calculates similarity between product vectors and user preferences using cosine similarity to provide relevant recommendations based on the highest similarity scores. The system collects user characteristic data including playing position, field type, material, and cost to form preference profiles, which are then matched with product data and connected to chatbot and search feature interfaces. The web application was developed using Bootstrap for the frontend and Flask for the backend, with the word2vec-based recommendation model integrated to display real-time recommendations. Functional testing through User Acceptance Testing (UAT) confirmed that the system effectively generates accurate soccer shoe recommendations according to user preferences.

REFERENCES

[1] K. Okholm Kryger, K. Mutamba, S. Mitchell, S. C. Miller, And S. Forrester, “Physical Performance And Perception Of Foot Discomfort During A Soccer-Specific Match Simulation. A Comparison Of Football Boots,” J. Sports Sci., Vol. 39, No. 9, Pp. 1046–1054, 2021, Doi: 10.1080/02640414.2020.1856461.

[2] E. Muslim, B. Brahmantyo, F. F. Utami, And V. B. Jasmine, “The Effect Of Football Cleats Studs Design To Amateur Football Player Performance Evaluation,” J. Eng. Manag. Ind. Syst., Vol. 7, No. 1, Pp. 1–6, 2019, Doi: 10.21776/Ub.Jemis.2019.007.01.1.

[3] C. Saha And D. Prasetya, “Sistem Rekomendasi Pada E-Commerce Menggunakan K-Nearest Neighbor,” Vol. 4, No. 3, Pp. 2355–7699, 2017.

[4] S. Kumari, “Job Recommendation System Using Nlp,” Int. J. Res. Appl. Sci. Eng. Technol., Vol. 11, No. 5, Pp. 2721–2728, May 2023, Doi: 10.22214/Ijraset.2023.52183.

[5] Didi Riswan, Heri Eko Rahmadi Putra, And Risfan Nazar Saputra, “Pengembangan Sistem Rekomendasi Berbasis Kecerdasan Buatan Untuk Meningkatkan Pengalaman Pengguna Di Platform E-Commerce,” J. Komput. Teknol. Inf. Dan Sist. Inf., Vol. 2, No. 3, Pp. 572–580, 2024, Doi: 10.62712/Juktisi.V2i3.145..

[6] M. Tamam Huda And A. Permana Wibowo, “Recommendation System For Mobile Application Tour Guide And Travel Services Using Demographic Filtering And Content-Based Filtering Methods Based On Android”, [Online]. Available: Https://Jws.Rivierapublishing.Id/Index.Php/Jws

[7] D. R. Pradana, S. Sa’adah, And D. Nurjanah, “Sistem Rekomendasi Sepatu Lokal Menggunakan Metode Collaborative Filtering Pada Toko Sepatu Tarsius Store,” E-Proceeding Eng., Vol. 9, No. 3, Pp. 2216–2176, 2022.

[8] A. Mulyana And S. Yuliyanti, “Aplikasi E-Commerce Dengan Sistem Rekomendasi Berbasis Collaborative Filtering Pada Toko Distro Nocturnal.”

[9] D. Radhian And I. Afrianto, “Pembangunan Aplikasi Chatbot Sebagai Media Pencarian Informasi Dalam Bidang Peternakan.”.

[10] R. F. Muharram Et Al., “Implementasi Artificial Intelligence Untuk Deteksi Masker Secara Realtime Dengan Tensorflow Dan Ssdmobilenet Berbasis Python,” J. Widya, Vol. 3, No. 2, Pp. 281–290, [Online]. Available: Https://Jurnal.Amikwidyaloka.Ac.Id/Index.Php/Awl.

[11] F. R. Hariri And L. W. Rochim, "Sistem Rekomendasi Produk Aplikasi Marketplace Berdasarkan Karakteristik Pembeli Menggunakan Metode User Based Collaborative Filtering," Teknika, Vol. 11, No. 3, Pp. 208–217, Nov. 2022, Doi: 10.34148/Teknika.V11i3.538.

[12] A. H. Ritdrix And P. W. Wirawan, "Sistem Rekomendasi Buku Menggunakan Metode Item-Based Collaborative Filtering," J. Masy. Inform., Vol. 9, No. 2, Pp. 24–32, 2018, Doi: 10.14710/Jmasif.9.2.31482.

[13] R. H. Mondi And A. Wijayanto, "Recommendation System With Content-Based Filtering Method For Culinary Tourism In Mangan Application.".

[14] Aditya Quantano Surbakti, Regiolina Hayami, And Januar Al Amien, "Analisa Tanggapan Terhadap Psbb Di Indonesia Dengan Algoritma Decision Tree Pada Twitter," J. Coscitech (Computer Sci. Inf. Technol., Vol. 2, No. 2, Pp. 91–97, 2021, Doi: 10.37859/Coscitech.V2i2.2851.

[15] S. J. Angelina, A. Bijaksana, P. Negara, And H. Muhardi, "Analisis Pengaruh Penerapan Stopword Removal Pada Performa Klasifikasi Sentimen Tweet Bahasa Indonesia," Juara (Jurnal Apl. Dan Ris. Inform., Vol. 02, No. 1, Pp. 165–173, 2023, Doi: 10.26418/Juara.V2i1.69680