Proceedings of the 10th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS'24)
Barcelona, Spain - August 19-21, 2024
Paper No. MVML 122
DOI: 10.11159/mvml24.122

Personalized Fashion Product Recommendations using Transfer Learning and Nearest Neighbors Models

Ikhlass Boukrouh[1], Faouzi Tayalati[2], Abdellah Azmani[3]
[1, 2, 3]Intelligent Automation and BioMedGenomics Laboratory
FST of Tangier, Abdelmalek Essaâdi University, Tetouan, Morocco
[1]ikhlass.boukrouh@etu.uae.ac.ma; [2]faouzi.tayalati@etu.uae.ac.ma; [3]a.azmani@uae.ac.ma

Abstract **-** In the fashion retail e-commerce sector, personalized product recommendations are crucial for enhancing the shopping experience. This study introduces a method that combines a pre-trained deep learning model named VGG19 with the 10 nearest neighbors algorithm to recommend visually similar products. VGG19 is utilized to extract detailed features from product images, enabling more accurate recommendations. The nearest neighbors algorithm then selects the ten products most similar to those previously viewed by customers. Recommendations are ranked based on customer purchase frequency to prioritize the most popular and relevant items. This method's practical applicability was demonstrated by testing it on a diverse set of products, including jackets from outerwear, baby bodysuits from children's wear, socks from footwear, and sunglasses from the accessories category.

Keywords**:** Recommendation system - Product recommendation - Transfer learning - VGG19 - Nearest neighbors algorithm

1. Introduction

The rapid growth of e-commerce has transformed the retail landscape, particularly in the fashion sector, where individuals are often assessed based on their clothing choices. In this dynamic environment, where customer satisfaction is essential, retailers are increasingly employing advanced technologies to enhance the shopping experience and drive engagement. Two key concepts, recommendation systems, and sentiment analysis, are emerging as essential applications for researchers [1]. Personalized product recommendations aim to guide consumers through the vast array of options available [2], ensuring they find items that resonate with their style and preferences.

Traditional recommendation systems often rely on collaborative filtering or content-based methods, whether in the realm of job seeking [3], suggesting movies [4], or recommending books [5]. While successful, they can face challenges in capturing the detailed preferences of fashion shoppers, especially as their preferences change. For instance, content-based methods, which rely on the purchase history of customers [6], may not adapt well to significant life changes, such as having a child, thus impacting the accuracy of recommendations. However, advancements in deep learning and computer vision offer promising solutions. By analysing product visual features, these techniques can provide more personalized recommendations, better aligning with individual preferences.

Additionally, recent studies have introduced innovative methods to enhance fashion recommendation systems, including the multi-task learning and gender-aware fashion recommendation system detailed in [7]. This system integrates gender and object detection, similarity generation, and recommendation components to improve efficiency and quality, outperforming existing approaches. Additionally, another approach proposed in [8] utilizes OpenCV for real-time video frame cloth detection, followed by feature extraction using a Convolutional Neural Network (CNN) model and K-Nearest Neighbors (K-NN) selection of similar items, with user feedback analysis to further refine recommendations.

This study contributes to the advancement of personalized product recommendation systems by introducing an innovative approach that combines deep learning techniques with traditional recommendation algorithms. The proposed methodology uses a pre-trained deep learning model for extracting features from product images. These features are then integrated with traditional recommendation algorithms to offer customers personalized product suggestions based on their browsing history.

The remainder of this paper is organized as follows: Section 2 presents our methodology, detailing the approach used for feature extraction and recommendation. Section 3 discusses the experimental setup and results obtained. Finally, Section 4 concludes the paper and suggests some future research directions.

2. Proposed methodology

The approach utilizes visual similarity to enhance product recommendations in the fashion retail e-commerce sector. As shown in Fig. 1, the methodology utilizes a CNN model, specifically the pre-trained VGG19, to extract features from product images. Initially, data is collected and pre-processed according to the requirements of the VGG19 model. Once the features are extracted, the K-NN algorithm is used to identify the most relevant products based on the input image, which is the product image consulted by the customer. The recommendations are then further ranked based on customer purchase frequency to ensure they are both visually relevant and popular.

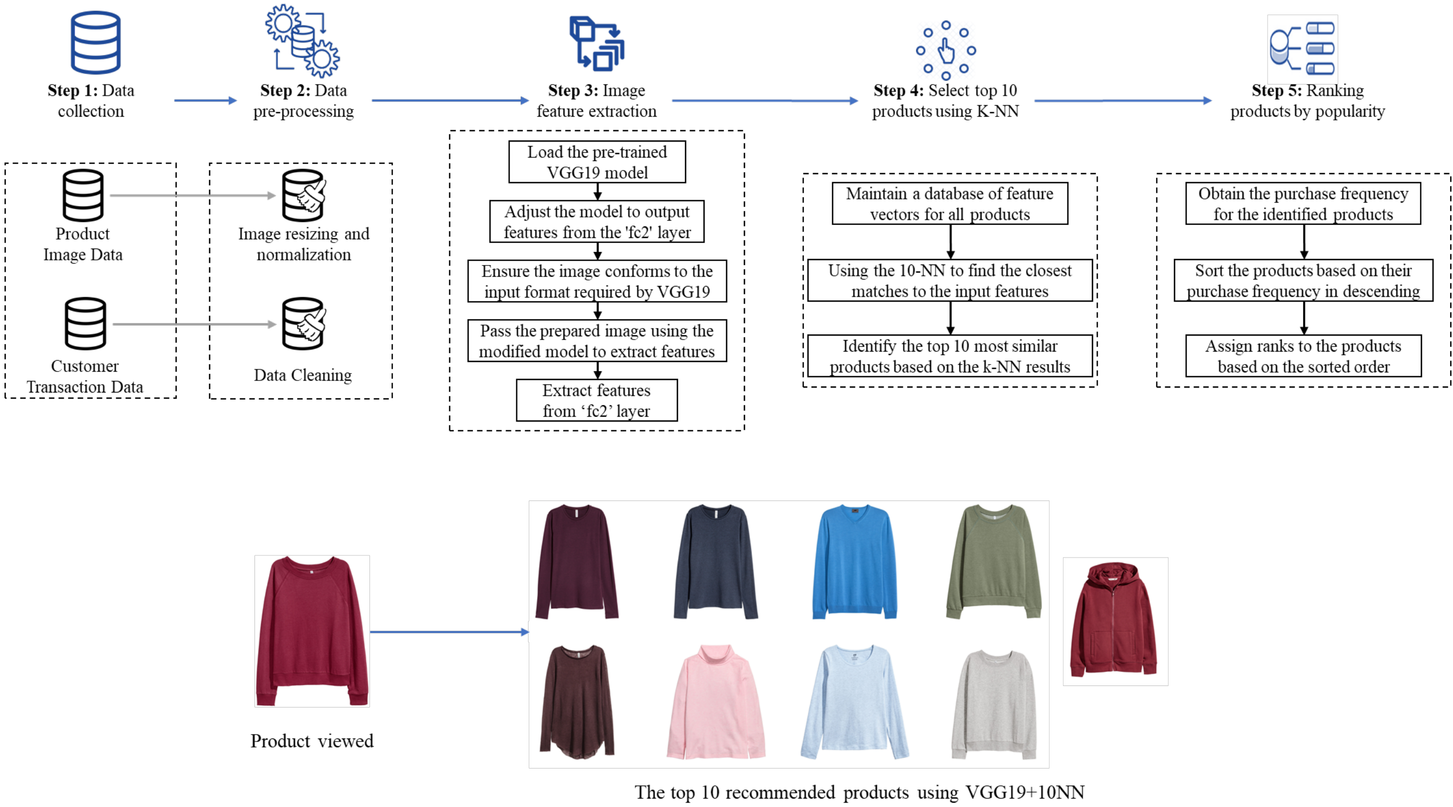

Fig. 1: Proposed approach for product recommendations.

2.1. Image Feature Extraction

Several pre-trained models, such as ResNet [9], MobileNet [10], and VGGNet [9], [11], are available for extracting high-level features from product images using CNN. In this study, VGG19 was chosen because it outperformed the other models in terms of feature extraction quality and similarity calculations [9]. This process ensures that the visual characteristics of the products are captured effectively for accurate similarity calculations and recommendations.

The VGG19 model, which stands for Visual Geometry Group 19-layer network, is a deep CNN architecture developed by the Visual Geometry Group at the University of Oxford [12]. VGG19 is renowned for its simplicity and effectiveness, consisting of 19 layers, including convolutional, pooling, and fully connected layers. This model has a total of 138 million parameters. VGG19 is pre-trained on a subset of the ImageNet dataset [13], which is used in the ImageNet Large-Scale Visual Recognition Challenge, containing over a million images, enabling it to classify images across thousands of classes [13]. This extensive training has equipped VGG19 with the ability to capture detailed and varied feature representations from a wide range of images, thereby making it highly effective in feature extraction tasks.

VGG19 comprises a series of layers [11], including 16 convolutional layers responsible for applying convolution operations to input images using multiple filters to detect features like edges, textures, and patterns. Additionally, it incorporates 19 learnable weights layers [11], utilized for transfer learning purposes. These layers include pooling layers, which down-sample feature maps to reduce spatial dimensions while retaining critical information, and fully connected (FC) (FC) layers. FC layers are dense, connecting every neuron in one layer to every neuron in the next, commonly used for classification tasks at the network's end. In this methodology, the 'fc2' layer, the second fully connected layer of VGG19, is is particularly important. It contains 4,096 neurons and captures high-level features of the input images. By extracting features from this layer, a rich and comprehensive representation of the image's visual characteristics is obtained. The output output of the model is a feature vector that serves as a detailed representation of the input image, which will be used for similarity measurement. The architecture of VGG19 used in this study is illustrated in Fig. 2.

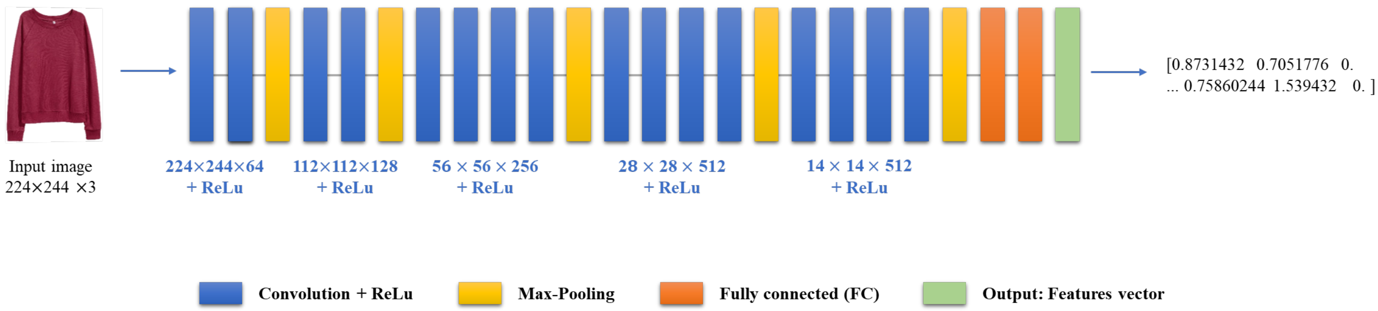

Fig. 2: Modified VGG19 model to extract the feature vector.

The first step in the image feature extraction process involves loading the VGG19 model, pre-trained on the ImageNet dataset, to leverage its pre-learned weights and architecture. The model is then adjusted to output features from the 'fc2' layer, providing a high-dimensional feature vector that serves as the basis for similarity comparison with other products. Once the model is prepared, the input image must be pre-processed to conform to the input requirements of VGG19. This involves resizing the image to 224x224 pixels, converting it to an array, and normalizing the pixel values. After pre-processing, the image is passed through the modified VGG19 model to extract the feature vector from the 'fc2' layer. The resulting output is a high-dimensional feature vector that captures the essential visual characteristics of the input image. This feature vector is crucial for the subsequent similarity comparison with other product images.

2.2. Product Recommendation

After image feature extraction, the next step is to identify visually similar products and rank them based on their popularity using the K-NN algorithm and similarity scores derived from feature vectors.

The K-NN algorithm is central to the product recommendation process and requires two key metrics [14]: the number of neighbors K and a distance metric to measure similarity. In this study, the distance metric used is cosine similarity, which measures the angle between two vectors, providing a value between 0 and 1. Higher values signify greater similarity between the vectors [5], with 1 indicating identical vectors.

The similarity between the input image's feature vector and the feature vectors of other products in the database is quantified using cosine similarity. The cosine similarity S between two vectors A and B given by Eq. (1) [5].

$$S(A,B) = \frac{A \cdot B}{\|A\| \|B\|} \tag{1}$$

Where: $A \cdot B$ is the dot product of vectors A and B, calculated as follows: $A \cdot B = a_1 b_1 + \cdots + a_n b_n$, where a_i and b_i are the components of the feature vectors of the input image and another product image, respectively. $\|A\|$ is the norm of vectors A, and calculated $\|A\| = \sqrt{a_1^2 + \cdots a_n^n}$.

Once the similarity scores are calculated, the K-NN algorithm selects the top similar products. Given an input feature vector F_{input} and a set of feature vectors F_i in the database, the top K products are identified by calculating the similarity score $S(F_{input}, F_i)$ for each product i in the database. The products are then sorted by their similarity scores, and the top products with the highest scores are selected for further analysis [14]. For this study, K is set to 10, meaning the top 10 most similar products are identified and prioritized.

After identifying the top 10 similar products using K-NN, these products are ranked based on their purchase This ensures that the recommendations are not only visually similar but also popular among customers. Given the purchase frequency P_i for each product i, the products are sorted in descending order based on these frequencies.

3. Experiment and Results

This section presents the empirical evaluation of the proposed product recommendation system. This section details the experimental setup, the datasets used, the data preprocessing, and the results obtained from the experiments. The goal is to demonstrate the effectiveness of the methodology in recommending visually similar and popular products.

The datasets used in this study were sourced from Kaggle [15], a popular platform for data science and machine learning competitions. The first dataset includes 105542 images of various clothing items, such as tops, pants, bottes, dresses, and accessories. Fig. 3 shows samples of product images from the dataset. The second dataset contains purchase information for 363798 customers, detailing the frequency and history of their purchases during the years 2018, 2019, and 2020. The purchase information dataset was pre-processed to handle missing values and resolve inconsistencies, ensuring data quality and relevance.

Fig. 3: Samples of product images from the dataset.

For the product image dataset, the images were pre-processed to conform to the input requirements of the VGG19 model. This involved resizing all images to 224x224 pixels to match the input size required by VGG19 and normalizing the pixel values to the range [0, 1] by dividing by 255. Fig. 4 illustrates an example of an image before resizing, after resizing to 224x224 pixels.

Fig. 4: Original and Resized Image (224x224 pixels).

The VGG19 model, pre-trained on ImageNet, was used to extract high-dimensional feature vectors from the prepared images. The model was adjusted to output features from the 'fc2' layer, providing detailed visual characteristics for similarity comparison. Using these feature vectors, the K-NN algorithm computed similarity scores and identified the top 10 similar products. Fig. 5a to 5d illustrate the recommendations for a jacket, baby bodysuits, socks, and sunglasses, respectively, showing the consulted product images along with their top recommended products and corresponding IDs. These products were then ranked based on their purchase frequency to ensure the recommendations were both visually relevant and popular.

Fig. 5a: Jacket recommendations for ID 0176209025.

Fig. 5b: Bodysuits recommendations for ID 0348330041.

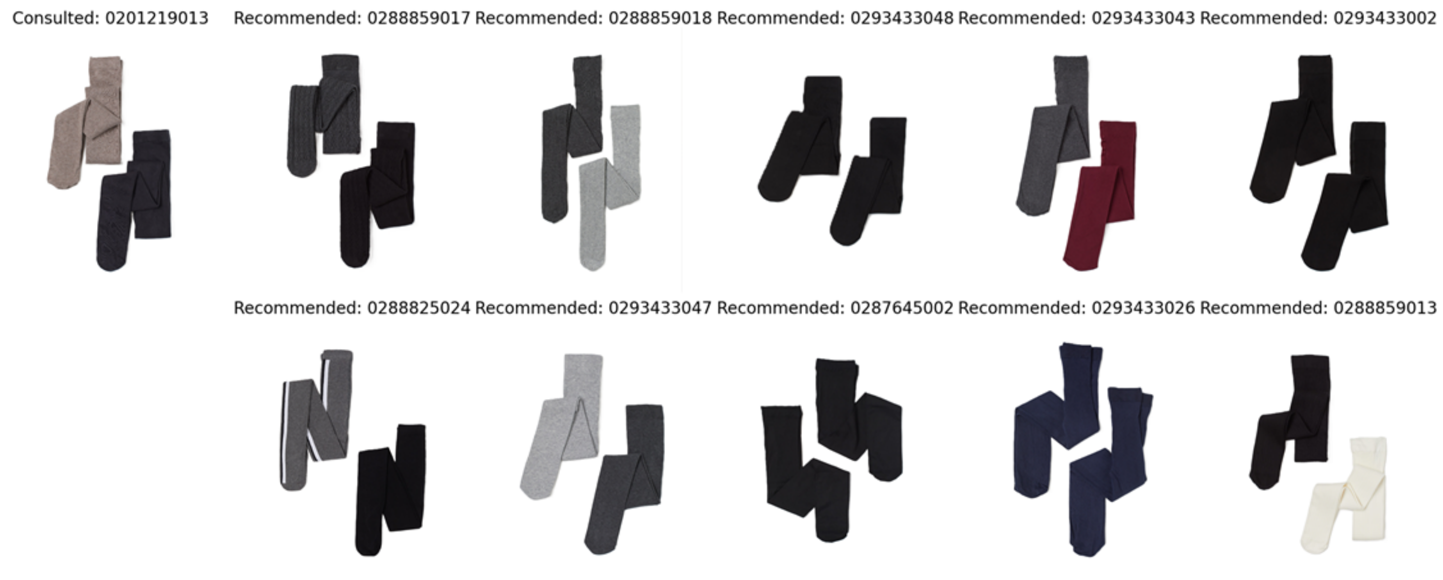

Fig. 5c: Socks recommendations for ID 0201219013.

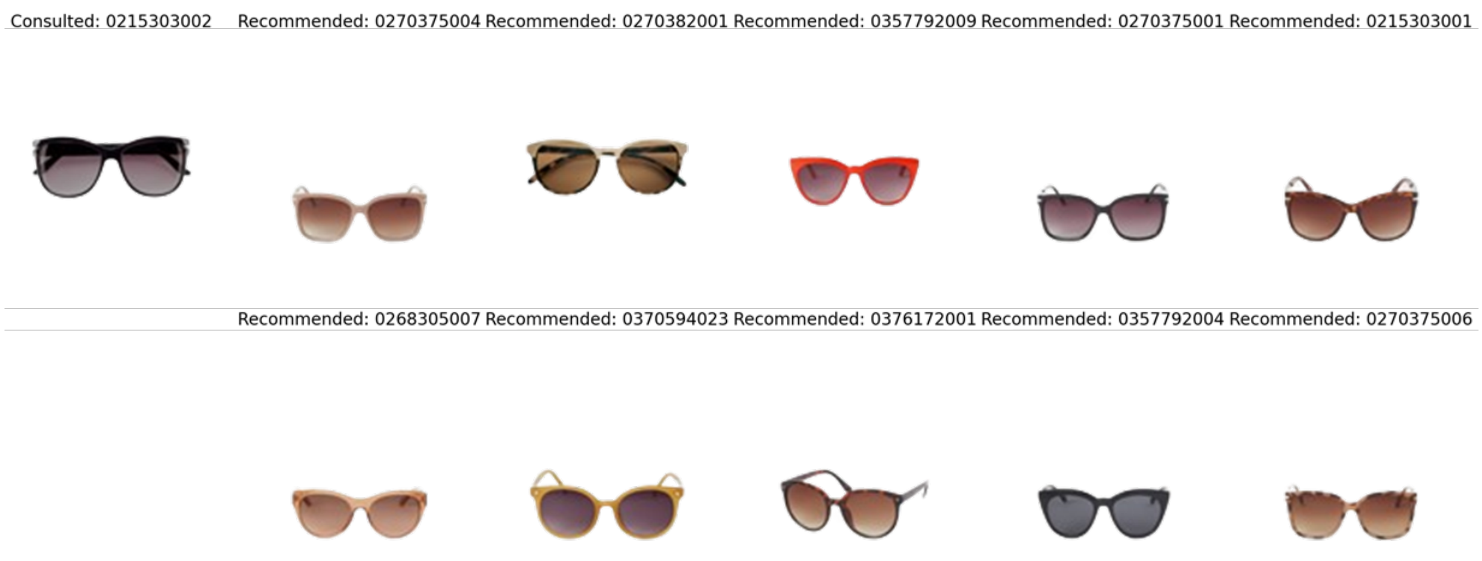

Fig. 5d: Sunglasses recommendations for ID 0215303002.

After selecting the top 10 similar products for each of the four consulted products, the purchase frequency for each recommended product was calculated. Table 1 presents the purchase frequencies for the recommended sunglasses, providing insight into the popularity of each item. Fig. 6 illustrates how the recommendations will appear to customers for the sunglasses. It is noted that one of the recommended sunglasses (ID: 0270375006) does not appear in the final recommendation because it has never been sold.

Table 1: Recommended sunglasses frequencies.

Recommended ID	Purchase frequency
0376172001	288
0215303001	270
0357792004	166
0270375001	114
0270382001	98
0370594023	78
0268305007	69
0357792009	58
0270375004	27
0270375006	0

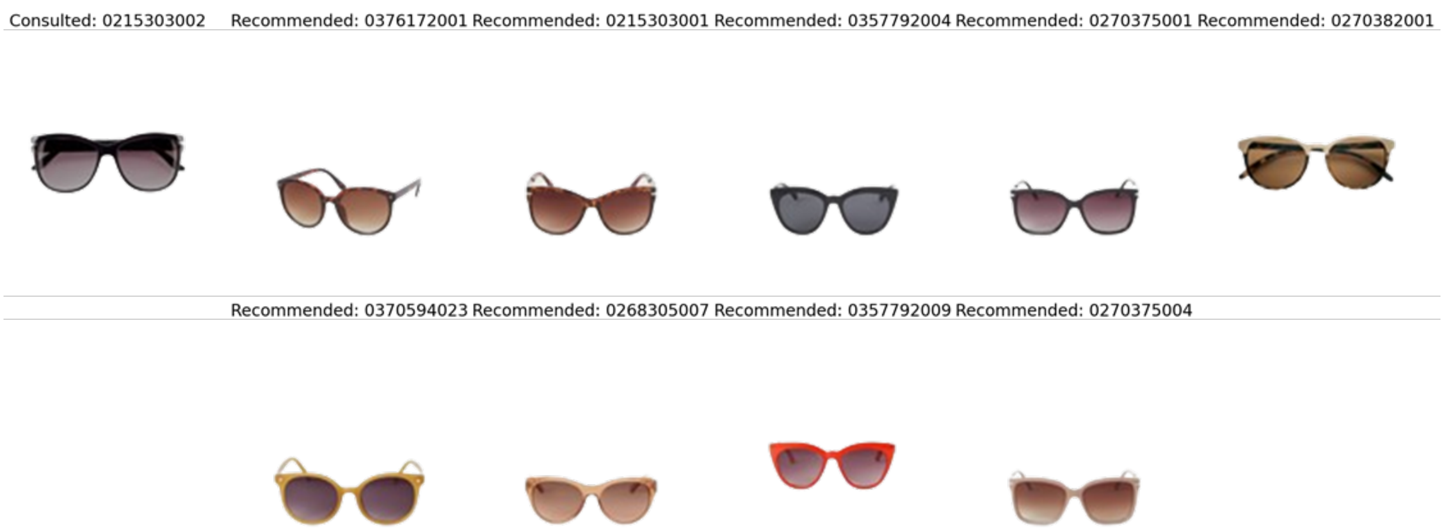

Fig. 6: Ranked Product Suggestions for ID 0215303002.

4. Conclusion

A visually oriented framework for fashion product recommendation is presented in this paper. By utilizing a pre-trained CNN, specifically VGG19, detailed feature vectors are extracted from product images. This approach involves two stages: extracting high-dimensional visual features from the images and applying the K-NN algorithm to identify and rank similar products based on these features. The system effectively recommends visually similar and popular products. Datasets from Kaggle, containing comprehensive collections of product images and purchase histories, were used. The purchase frequency for each recommended product was calculated, providing insights into their popularity and ensuring that the recommendations are visually relevant and widely accepted.

Future work could explore integrating real-time user feedback to continually refine the accuracy of recommendations. Additionally, incorporating dynamic user behavior data could further improve the system's performance and applicability.

Acknowledgements

The support for this research is provided by the Ministry of Higher Education, Scientific Research, and Innovation, as well as the Digital Development Agency (DDA) and the National Center for Scientific and Technical Research (CNRST) of Morocco, under the Smart DLSP Project - AL KHAWARIZMI AI PROGRAM.

References

[1] I. Boukrouh and A. Azmani, 'ARTIFICIAL INTELLIGENCE APPLICATIONS IN E-COMMERCE: A BIBLIOMETRIC STUDY FROM 1995 TO 2023 USING MERGED DATA SOURCES', *Int. J. Prof. Bus. Rev.*, vol. 9, no. 4, p. e4537, Apr. 2024, doi: 10.26668/businessreview/2024.v9i4.4537.

[2] M. Sridevi, N. ManikyaArun, M. Sheshikala, and E. Sudarshan, 'Personalized fashion recommender system with image based neural networks', *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 981, p. 022073, Dec. 2020, doi: 10.1088/1757-899X/981/2/022073.

[3] S. Eliyas and P. Ranjana, 'Recommendation Systems: Content-Based Filtering vs Collaborative Filtering', in *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Greater Noida, India: IEEE, Apr. 2022, pp. 1360–1365. doi: 10.1109/ICACITE53722.2022.9823730.

[4] H. Fulzele, M. Bhoite, P. Kanfade, A. Yadav, M. Sahu, and A. Thomas, 'Movie Recommender System using Content Based and Collaborative Filtering', *Int. J. Innov. Sci. Res. Technol.*, vol. 8, 2023.

[5] V. Nuipian and J. Chuaykhun, 'Book Recommendation System based on Course Descriptions using Cosine Similarity', in *Proceedings of the 2023 7th International Conference on Natural Language Processing and Information Retrieval*, Seoul Republic of Korea: ACM, Dec. 2023, pp. 273–277. doi: 10.1145/3639233.3639335.

[6] K. K. Yadav, H. K. Soni, and N. Pathik, 'Recommendation System Based on Double Ensemble Models using KNN-MF', *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 5, 2023, doi: 10.14569/IJACSA.2023.0140566.

[7] A.-Z. Naham, J. Wang, and A.-S. Raeed, 'Multi-Task Learning and Gender-Aware Fashion Recommendation System Using Deep Learning', *Electronics*, vol. 12, no. 16, p. 3396, Aug. 2023, doi: 10.3390/electronics12163396.

[8] D. Rathod, A. Tundare, P. Shelke, S. Muneshwar, and S. L. Dawkhar, 'FASHION RECOMMENDATION USING DEEP LEARNING', *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 06, no. 01, Jan. 2024.

[9] R. Mostafiz, M. Rahman, A. Islam, and S. Belkasim, 'Focal Liver Lesion Detection in Ultrasound Image Using Deep Feature Fusions and Super Resolution', *Mach. Learn. Knowl. Extr.*, vol. 2, no. 3, pp. 172–191, Jul. 2020, doi: 10.3390/make2030010.

[10] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam., 'MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications'. arXiv, 2017. doi: 10.48550/ARXIV.1704.04861.

[11] Rashid, M., Khan, M. A., Alhaisoni, M., Wang, S. H., Naqvi, S. R., Rehman, A., & Saba, T, 'A Sustainable Deep Learning Framework for Object Recognition Using Multi-Layers Deep Features Fusion and Selection', *Sustainability*, vol. 12, no. 12, p. 5037, Jun. 2020, doi: 10.3390/su12125037.

[12] K. Simonyan and A. Zisserman, 'Very Deep Convolutional Networks for Large-Scale Image Recognition'. arXiv, 2014. doi: 10.48550/ARXIV.1409.1556.

[13] Y. Zheng, C. Yang, and A. Merkulov, 'Breast cancer screening using convolutional neural network and follow-up digital mammography', in *Computational Imaging III*, A. Ashok, J. C. Petruccelli, A. Mahalanobis, and L. Tian, Eds., Orlando, United States: SPIE, May 2018, p. 4. doi: 10.1117/12.2304564.

[14] D. R. Anamisa, A. Jauhari, and F. Ayu Mufarroha, 'K-Nearest Neighbors Method for Recommendation System in Bangkalan's Tourism', *ComTech Comput. Math. Eng. Appl.*, vol. 14, no. 1, pp. 33–44, May 2023, doi: 10.21512/comtech.v14i1.7993.

[15] C. García Ling, ElizabethHMGroup, J. Ferrando, and FridaRim, 'H&M Personalized Fashion Recommendations.' 2022. [Online]. Available: https://kaggle.com/competitions/h-and-m-personalized-fashion-recommendations