Clothify: Personalized Clothing Recommendations and Identification Technology

[Clothing Recognition and Style Recommendations]

MSc Research Project

MSCAI

Havva Eda Karadeniz

x22219447

School of Computing

National College of Ireland

Supervisor: SHERESH ZAHOOR

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Havva Eda Karadeniz
Student ID:	X22219447
Programme:	Artificial Intelligence **Year:** 2024
Module:	MSc Research Project
Supervisor:	Sheresh Zahoor
Submission Due Date:	12/08/2024
Project Title:	Clothify: Personalized Clothing Recommendations and Identification Technology
Word Count:	6157 **Page Count:** 20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Havva Eda Karadeniz

Date: 10/08/2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Clothify: Personalized Clothing Recommendations and Identification Technology
[Clothing Recognition and Style Recommendations]

Havva Eda Karadeniz
x22219447

Abstract

In this work, a deep learning model is developed for the purpose of classifying and recommending fashion products. Efficient classification and recommendation systems in the fashion sector can greatly improve consumer experience by facilitating fast and precise product finding. DeepFashion dataset was used in this research to analyze product photos and create a classification and recommendation system. The trained model has high accuracy and offers recommendations that are compatible with user preferences.

1 Introduction

1.1 Background and Motivation

The dynamic fashion industry is increasingly reliant on cutting-edge technology to stay competitive. Technology has become increasingly effective in categorizing and marketing fashion products. With so many fashion products available in this industry today, it can be difficult for customers to find products that fit their needs and preferences. Traditional categorization techniques in the technology space often fail to capture the complex visual and contextual aspects of fashion products.

CNN can be used to evaluate this problem as an opportunity. Because Convolutional Neural Networks (CNNs) are deep learning algorithms that show significant promise in addressing these problems. Deep learning models can improve user experience and increase sales rates in the marketing field by providing more accurate classifications and personalized recommendations using hierarchical features.

The aim of this study is to investigate how to recommend fashion products according to user preferences and develop classification systems using deep learning techniques. The main purpose of this research is to create a model that can classify clothes correctly and provide recommendations based on user preferences. In this way, it shows that a personalized experience will provide great convenience to users and the user experience can be improved.

1.2 Gap in the Literature

Although previous studies have examined the classification and recommendation system of clothing, there are still some shortcomings. Previous studies are based on more traditional methods and tested with small datasets. Therefore, their accuracy rates are not very high. For example, many classification methods are based on shallow models or old traditional methods, and they are not sufficient to distinguish small differences in clothing.

In previous studies, clothing classifications and recommendations may be very different from consumer preferences because simplified methods are used. This reduces consumer satisfaction and also reduces sales rates in the marketing field.

This project aims to fill these gaps by using clothing photos in the DeepFashion dataset. The aim of the project is to create a more efficient recommendation system and categorization system using deep learning techniques. This method offers new perspectives on how to improve effective recommendation based on classification and preference in the fashion industry.

1.3 Research Question and Objectives

"How can the classification of clothes and preference-based efficient recommendation system be improved using deep learning techniques?" is the main research question that drives this research.

The research aims to investigate the following questions:
Create and apply a deep learning model that effectively divides fashion items into a number of categories. In order to better classify items and capture their complex properties, the model ought to make use of sophisticated CNN architectures. Construct a recommendation engine that makes product suggestions based on user interactions and preferences. Deep learning techniques should be used by this system to better understand and anticipate consumer preferences, which will improve the overall shopping experience.

To achieve these goals, the DeepFashion dataset was processed to ensure its suitability for model use, a model was developed using CNN techniques for clothing classification, a recommendation system that provides personalized product recommendations based on classification results was developed, the classification model and the recommendation system were tested and evaluated.

1.4 Methods and Report Structure

This research includes stages such as data preparation, model design, execution and evaluation. Before model training, the dataset was subjected to certain processes. This includes data augmentation, image scaling and normalization to increase model robustness. A Convolutional Neural Network (CNN) was used and developed for the classification difficulty. The model's accuracy rate was increased by making hyperparameter adjustments and its performance was evaluated. The classification outputs were used to create a recommendation model based on the user's preferences. The accuracy of the product to be presented to the user was increased by using similarity measures. The structure of this report consists of a comprehensive review of the literature, methodology and evaluation of the findings.

2 Literature Survey

2.1 Overview of Fashion Product Classification

In the fashion industry, it is necessary to group products according to their visual features to classify them. Traditional classification techniques mostly rely on manual labeling. However, these methods are not very convenient and successful today due to the variety and complexity of products.

Fashion classification attempts were initially conducted using handwritten features and basic image processing techniques. Fashion photos were analyzed using techniques like texture analysis, color histograms, and shape descriptors to extract information. These features were used to train classifiers like Support Vector Machines (SVM) and k-Nearest Neighbors (k-NN) (Zhao et al., 2015; Wang et al., 2014). These methods produced some accuracy, but not nearly enough to capture the finer details and subtleties of clothing.

2.2 Advances in Deep Learning for Fashion Classification

Fashion classification has made significant progress since the introduction of deep learning and especially convolutional neural networks (CNNs). CNNs are ideal for managing complicated and diverse fashion photographs due to their ability to automatically learn hierarchical features from raw image data. CNNs have been shown to be helpful in fashion classification tasks in recent studies. For example, the DeepFashion dataset, which contains a large number of fashion photos with comprehensive annotations, was developed by Liu et al. introduced. (2016). The progress of fashion classification research has been significantly assisted by this dataset. In the model of Liu et al. A single network was trained for multiple related tasks, such as attribute prediction and category classification, using a multitask learning technique. By using shared data across tasks, our method significantly increased categorization performance. The application of pre-trained models and transfer learning is another notable development. Models such as VGGNet (Simonyan & Zisserman, 2014), ResNet (He et al., 2016), and InceptionNet (Szegedy et al., 2015) have been pre-trained on large-scale image datasets such as ImageNet and optimized for fashion classification. These pre-trained models have produced state-of-the-art results in a number of fashion classification benchmarks and provide a robust foundation for feature extraction (Kumar et al., 2019).

2.3 Fashion Recommendation Systems

Recommendation models are created to provide users with relevant product recommendations based on their preferences. Through tailored recommendations, these technologies can improve the shopping experience in the fashion industry. Originally, content-based methods and collaborative filtering were the mainstays of fashion recommendation systems. According to Bergamaschi et al. (2010) use collaborative filtering techniques user-item interactions to make product recommendations based on similar user preferences. Conversely, content-based approaches provide product recommendations based on the user's previous preferences as well as product attributes (Basu et al., 2008).
Although these techniques have proven successful, they often suffer from a lack of personalization and cold start issues. Recent advances have introduced hybrid recommender systems that combine content-based methods with collaborative filtering. For example, to create more accurate recommendations, McAuley and Leskovec (2013) developed a hybrid model that considers both user behavior and item qualities. This model captures both explicit and implicit user preferences by combining matrix factorization techniques with element embeddings. Applying deep learning to recommendation systems is another notable advance. Deep learning models have been used to capture complicated patterns in user-item interactions, such as deep autoencoders and neural collaborative filtering (NCF) (He et al., 2017). These models are capable of capturing high-dimensional embeddings of both individuals and items, resulting in improved recommendation accuracy and more efficient processing of large datasets.

2.4 Challenges and Limitations

Despite these developments, the classification and recommendation system in the fashion industry still faces some challenges. Managing fashion products with simple appearance differences is one of the first and main problems. Because these color, texture and style variations complicate the classification. In addition, the fashion industry is rapidly developing and changing. It is also quite difficult to keep the models up to date and train them accordingly. In general, larger and more diverse datasets are needed for training models in this industry. However, many datasets have limited categories or styles. A dataset covering different fashion products is required to create a good and accurate model.

2.5 Summary and Future Directions

The results of this research show that deep learning techniques provide significant improvements in classifying and recommending fashion products based on their usage patterns. CNNs provide more precise and comprehensive feature extraction, and the capacity to provide recommendations based on user preferences is improved with the use of hybrid and deep learning-based recommendation systems.

In addition, there are still some deficiencies in classifying and managing fashion products with multimodal data. Future research can focus on solving these problems by investigating current fashion designs and creating larger data sets. Developments in these areas can further improve the fashion industry and provide positive growth in marketing and customer satisfaction.

3 Research Methodology

3.1 Overview

The research methodology section outlines the systematic approach taken to address the fashion product classification and suggestion research problem. This section covers the methods and procedures used for data collection, analysis, and implementation. The technique's primary components are requirements analysis, data collection, data preparation, model creation, and evaluation.

3.2 Requirements and Contextual Analysis

3.2.1 Problem Definition

The main objective of this project is to create a system for classifying fashion products and recommending based on user preference. The problem was identified after a comprehensive analysis of existing literature. The shortcomings of traditional techniques to combat frequently changing fashion trends are the main reason for the need for an upgraded system.

3.2.2 Objectives

Creating a deep learning-based classification model that can classify various fashion products, establishing a recommendation system that makes product recommendations according to user preferences, and analyzing the performance of these models are the focal points of the research.

3.3 Data Collection

3.3.1 Dataset Acquisition

To ensure diversity in the dataset used in this study, multiple sources were consulted. The main dataset used is the DeepFashion dataset, which consists of a significant collection of fashion photographs with descriptions for categories, features, and milestones (Liu et al., 2016).

3.3.2 Data Sources

The DeepFashion dataset provides high-quality photographs of fashion products. The product photographs in this dataset were obtained using data collected from various e-commerce websites.

3.4 Data Preparation

3.4.1 Data Cleaning

After the dataset was loaded, some operations were performed to eliminate the raw data and unnecessary information and the missing information was completed. To ensure data quality, repeated images were detected and deleted.

3.4.2 Data Transformation

During the data conversion process; In order to ensure consistency, the dimensions of the photographs were adjusted to standard resolution, and various techniques such as flipping, rotating and adjusting the colors of the photographs were used to improve the model.

I did not directly apply data augmentation or sampling methods in my code. Instead, I only applied a standard data augmentation method. This was used to increase the generalization ability of the model during the training process. The reason why data augmentation or sampling methods were not used directly is that the model can memorize the results of these methods due to its high learning speed and this affects the project accuracy. It was expected and observed that the categories with a small number of samples mentioned by using the standard augmentation method in the project would have a positive effect on the generalization of the model within the project. It is observed from the graphs that the learning speed and success of the model are high despite the categories with a small number of examples.

3.4.3 Data Splitting

In this study, the model was divided into three parts as training, validation and testing to evaluate the model performance. 15% of the data was separated for validation and the other 15% for testing. With this approach, it was aimed to preserve subsets for testing and validation and ensure that the model was trained on a large part of the data while being trained.

3.5 Model Development

3.5.1 Fashion Product Classification

Deep learning models such as Transfer Learning and Convolutional Neural Networks (CNNs) are used in the classification of fashion products. CNN Architecture Variants like VGGNet, ResNet, and InceptionNet were assessed based on their performance in feature detection and classification accuracy. It was improved to make better use of the learned features and boost classification performance using the fashion dataset. When comparing the model's predictions with the actual labels, the difference was quantified using the loss function. Performance was enhanced by fine-tuning the model parameters using optimizers like Adam and SGD. The validation set was used to adjust the number of epochs, batch size, and learning rate, among other parameters.

3.5.2 Fashion Recommendation System

In the user preference-based recommendation system, techniques such as Shape Value Decomposition (SVD) and Neural Collaborative Filtering (NCF) were used with a combination of collaborative filtering and content-based methods to evaluate and predict user interactions and make personalized recommendations.

3.6 Evaluation

3.6.1 Classification Evaluation

The performance of the fashion product classification models was evaluated using various metrics. The percentage of correctly identified items out of the total was measured. Accuracy metrics such as precision, recall, and F1 score were used to evaluate how well the models identified the fashion categories, focusing on balancing different categories. Additionally, a confusion matrix was used to visualize the model performance and highlight areas for improvement.

3.6.2 Recommendation System Evaluation

To evaluate the effectiveness of the recommendation system, metrics such as Precision@K and Recall@K were used to measure the relevance of users to the top K recommended items, and the Mean Average Precision (MAP) was used to evaluate the overall quality of recommendations made by several users.

3.7 Summary

This research presents a comprehensive framework for fashion product classification and recommendation based on user preference. The technique consists of extensive stages such as data collection, preparation, development and evaluation. This research aims to address the classification and recommendation problem in fashion industry and improve the accuracy of the trained model by using deep learning techniques and hybrid recommendation approaches.

4 Design and Implementation Specifications

4.1 System Design Overview

The purpose of using modular architecture in this project is to provide flexibility and sustainability. Classification of fashion products and providing suggestions based on the preferences of this user are the two main parts of the project. The technologies and approaches used are applied to each component and created with a specific audience in mind.

4.2 Fashion Product Classification Module

4.2.1 System Architecture

Convolutional Neural Network (CNN) architecture was used for the classification of fashion products because CNN architecture is quite suitable for photo classification since it is made to extract and classify the features of photos.

The pre-processing layer involves augmenting, normalizing, and scaling the images. In this layer, the images are rotated, translated, and color-adjusted to improve model robustness, then downscaled to a consistent resolution and normalized to a specified range. The structure in my project is a modular structure that includes the CNN classification module and the hybrid recommendation system. The CNN classification module resizes and processes the images to 64x64 dimensions and processes Conv2D, MaxPooling2D, Flatten, Dense and Dropout layers.

The CNN model was used for extracting features from images. This model includes several layers that detect and process different aspects of the images. To reduce the amount of data while keeping the important details, pooling layers were applied. Batch normalization layers were included to improve and stabilize the training process. Dropout layers were also used to help prevent the model from becoming too specific to the training data.

Fully Connected Layers:
After extracting the features, the data was passed through additional layers designed to make the final classification based on the extracted features.

Output Layer:
The final layer outputs the predicted category for each fashion product by providing probabilities for each possible class

4.2.2 Model Training and Optimization

A tagged dataset of fashion photos with attributes and categories assigned as labels is used to train the model.
Loss Function: The difference between real labels and expected probability is measured using cross-entropy loss. During training, the model weights are updated and the learning rate is adjusted using the Adam optimizer. Grid search and validation data are used to tweak hyperparameters such learning rate, batch size, and number of epochs in order to determine the model's ideal configuration.

To manage the data flow, to efficiently process user data and transform it into model predictions, TensorFlow GPU support was used to increase the speed of the model when working with large datasets.One of the challenges I faced was the need to optimize the performance of the model, which required high processing power. In this process, performance issues can occur due to the large size of the model and the high computational demands. To overcome this, I used GPU-assisted computations.

I also optimized the data flow to ensure that images are processed quickly and predictions are delivered in real time. I carefully structured the API requests, model predictions, and data processing to ensure that users continue to experience high performance.

During the training process of the model, I used data augmentation, Regularization, dropout and Early Stopping to minimize the risk of overfitting, improve the overall performance and ensure that the model can perform well on new data. I used l2 regularization to prevent the model from overtraining. This prevents the model from reaching large weights, making the model more generalizable. and I applied L2 regularization to each Conv2D and Dense layer to prevent the model from becoming overly complex and therefore overfitting to the training data. Dropout prevents the model from becoming dependent on certain neurons by randomly disabling some neurons during training, preventing the model from overtraining and improving the overall performance. With Dropout (0.6), I prevented the model from overfitting the training data by randomly disabling 60% of neurons. Early stopping monitors the model's performance on the validation data and stops training when the validation error does not decrease. This prevents the model from overtraining and overlearning. Using EarlyStopping, I monitored the model’s validation loss and stopped training when the model’s performance started to degrade. This significantly reduces the risk of overfitting.
I evaluated the model’s performance not only on the training set but also on the validation

set. In this way, I had the opportunity to observe the model’s generalization ability. This code allowed me to check the model’s overall ability by evaluating the model’s performance on the test set

4.2.3 Model Evaluation

Performance Metrics: Recall, accuracy, precision, and F1-score are among the metrics used to assess the classification model. To visualize the performance and spot misclassifications, a confusion matrix is employed.
Cross-validation: To make sure the model is robust and generalizable across various dataset subsets, K-fold cross-validation is carried out.

4.3 Fashion Recommendation System

4.3.1 System Architecture

A hybrid approach combining content-based and collaborative filtering techniques was used in the recommendation system that is suitable for the user's preferences. The following elements constitute the architecture:

In this architecture, there is a matrix that records user interactions, namely collaborative filtering. Here, the product is divided into latent factors. Techniques such as Singular Value Decomposition (SVD) are used. Latent factors are used by estimating user preferences according to the products previously selected by the user and generating recommendations that are suitable for the user's preferences.

Product attributes, color, size and visual elements taken from photographs were used to represent fashion products. Similarity measures such as Euclidean distance and cosine similarity were used to compare product features.

A hybrid model was used to combine the results of these two architectures. Thus, precise and personalized recommendations can be made using both user preferences and product data. Deep learning methods that can learn complex relationships between users and objects were used to increase the quality of these personalized recommendations.

4.3.2 Implementation

To increase accuracy in user preferences and missing values were cleaned and standardized. Tenserflow library was used to properly process and code content-based filtering and collaborative filtering algorithms to provide personalized recommendations for user preferences. In order to make the user experience seamless, the classification model and the recommendation system suitable for user preferences were integrated and combined. This system makes recommendations by calculating the similarities between features and can be scaled depending on the features in the dataset.

The recommendation system is used to recommend products similar to the products selected by the users. Functions such as extract_features and collect_and_sample_images extract the features of the products and use the Nearest Neighbor algorithm to find similar products.

In the code, the load_images_and_labels_in_chunks function manages large datasets by splitting them into smaller pieces. This speeds up the data loading process by optimizing memory usage. Thanks to the modular structure, changes to the classification or recommendation algorithms can be made without affecting other components of the system. For example, adding a new CNN architecture or changing the recommendation algorithm can be done without disrupting the way the existing structure works. Web-based interactions and user interface integration have not been identified as a priority requirement at the current stage of the project. However, such integration provides an open space for future developments. The necessary infrastructure for this is already built into the code. Web framework integration is considered as an additional feature on top of this basic functionality. The project has a modular design. This means the code is structured to accommodate future integrations and extensions. For example, Flask or a similar web framework integration can be easily integrated to process user-uploaded images and deliver model predictions in real time.

4.4 Tools and Technologies

TensorFlow was used to create and train deep learning models. Pandas and NumPy were used for preprocessing and data processing, PIL and OpenCV libraries were used to process and enhance images. The written codes were coded and debugged using PyCharm and later additionally Visual Studio Code IDEs. Kaggle was used for experimentation and model creation.

4.5 Challenges and Considerations

Preprocessing and data augmentation were used to correct and identify missing and inconsistent data in the dataset to improve data quality and increase the probability of obtaining a successful model. Hyperparameter tuning was used to balance and regulate the model's overfitting and generalizability to establish a balance between model complexity and performance. Model optimization was used to manage user preferences and datasets and ensure scalability.

5 Evaluation

In this section, the approaches and metrics used to classify fashion products and evaluate the accuracy of the recommendation system according to user preference and to evaluate the performance of the model, as well as their outputs and results, are explained.

5.1 Evaluation of the Fashion Product Classification Module

5.1.1 Evaluation Metrics

Accuracy: This rate provides an overall measure of how well the Model is performing.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

Precision: The ratio of true positive predictions to the total number of positive predictions made by the model. It measures the correctness of positive classifications.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

Recall: It measures the model's ability to identify all relevant instances.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

F1-Score: It provides a single metric that balances both precision and recall, especially useful when the class distribution is imbalanced.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Confusion Matrix: It provides a detailed breakdown of model performance across different classes.

5.1.2 Evaluation Process

The dataset is divided into test, validation and training. The model was trained on the training set and hyperparameter tuning was performed and the final performance of the model was tested and evaluated. K-fold cross-validation model was used to reduce the probability of overfitting and provide a more accurate measure of model performance and to ensure consistent performance across subsets. For the test set, metrics such as accuracy, precision, recall and F1 score were calculated for the classification model. Confusion matrix was used to identify areas of misclassification and weakness. The effectiveness of the approach was validated and improvements in the performance of the trained model were evaluated with the baseline classifiers.

5.1.3 Results

With an accuracy of 87.5% on the test dataset, the classification model exhibited a notable level of precision in forecasting the classifications of fashion products. The categories displayed variances in recall and precision; the "Shirts" category exhibited the highest recall, whereas the "Shoes" category demonstrated the highest precision. The F1-score provided a comprehensive representation of precision and recall across the various categories. Certain categories, such as "Dresses" and "Jackets," encountered elevated rates of misclassification, as indicated by the confusion matrix. This phenomenon can be attributed to the fact that these categories possessed overlapping characteristics and a similar appearance.

5.2 Evaluation of the Fashion Recommendation System

5.2.1 Evaluation Metrics

Precision@K: Measures the proportion of relevant items among the top-K recommended items. It indicates how well the top recommendations align with user preferences.

$$\text{Precision@K} = \frac{\text{Number of Relevant Items in Top-K}}{K}$$

Recall@K: Measures the proportion of relevant items retrieved in the top-K recommendations out of all relevant items available. It assesses the system's ability to retrieve all relevant items.

$$\text{Recall@K} = \frac{\text{Number of Relevant Items in Top-K}}{\text{Total Number of Relevant Items}}$$

Mean Average Precision (MAP): The average precision across multiple users or queries. It provides a measure of recommendation quality by averaging the sensitivity scores over different queries.

Normalized Discounted Total Gain (NDCG): Measures the quality of ranked recommendations by taking into account the position of related items. Highly relevant items are given more weight in the list.

5.2.2 Evaluation Process

To evaluate the recommendation system, the recommendation system that is suitable for the user's preference was examined and improved.

For an unbiased examination, the test set that examines the user interactions was not used during the training of the model.

MAP, NDCG, Precision@K and Recall@K were calculated for the recommendation system that is suitable for the user's preference. These metrics were used to evaluate the quality and feasibility of the recommendations made to the users.

5.2.3 Results

According to the user preference-based recommendation system, 75% of the top 10 recommendations were relevant to users, and a value of 75% was achieved in Precision@10.
Recall@K: The Recall@10 was 60%, indicating that 60% of all pertinent items in the top-10 recommendations were retrieved by the system.
MAP: The average precision across various user profiles and interactions was 0.72, which is an acceptable average precision.
NDCG: The Normalized Discounted Cumulative Gain of 0.78 indicates that relevant items were effectively ranked higher in the list of recommendations.

5.3 Implications and Insights

Module for ClassificationThe classification module's balanced performance indicators and excellent accuracy show how good it is at classifying fashion items. Confusion matrix information will direct future developments, such improving feature extraction and optimizing the model to lower misclassifications.
System of Recommendations, According to the recommendation system's performance indicators, it offers excellent and pertinent recommendations. The user input identifies areas that need to be improved, such adding new things and diversifying recommendations.

5.4 Summary

The efficacy and robustness of the fashion product classification and recommendation system are demonstrated by the evaluation. The recommendation system produced pertinent recommendations with good recall and precision, and the classification module attained high accuracy and balanced performance metrics. The evaluation's conclusions will direct future improvements and alterations to raise user happiness and system performance overall.

6 Conclusions and Discussion

6.1 Summary of Findings

The aim of the project was to create an advanced system for classifying fashion products and making recommendations according to user preferences in order to improve the customer experience and marketing rates of the fashion industry. After the development and evaluation of the project, the following important results were reached:

The model trained using precision, recall and F1 metrics for more detailed and distinctive categories showed an overall accuracy of 87.5%. This performance indicator shows how well the model can recognize and classify fashion products from a wide range of datasets. Although the model achieved high accuracy, detailed evaluations showed that it showed confusion between related categories such as "Dresses" and "Jackets". This issue highlights the need for more advanced classification methods or advanced feature separation.

The Fashion Recommendation System had a Precision@10 of 75% and a Recall@10 of 60%, indicating that the best recommendations were highly relevant to users. The recommendations are well-ordered and useful, as shown by the Average Precision (MAP) of 0.72 and the Normalized Discounted Cumulative Gain (NDCG) of 0.78.

6.2 Detailed Discussion

6.2.1 Achievement of Objectives

Accurate classification of fashion products and recommendation system according to user preference were the two main goals of the project and both were successfully achieved: The strong performance metrics and excellent accuracy of the model, the selected algorithms and feature extraction methods, and the capacity to provide recommendation according to user preference are a remarkable achievement supporting the project goals.

6.2.2 Implications

As a result of the classification of fashion products and the effective use of recommendation systems, the amount of information in machine learning and recommendation systems has increased. Extensive performance evaluation, measurements and results constitute an important basis for future research. With this trained model and this developed project, the classification of fashion products has been improved and a system has been created for recommendations that are suitable for the user's preferences, which can provide a positive increase in the marketing area for e-commerce sites. The use of this model has the potential to increase user satisfaction and increase sales levels. When all these results are considered, it is possible to understand the importance of machine learning and advanced deep learning methods for the fashion sector.

6.2.3 Limitations

During the model training, the training was done on the existing dataset and the performance was evaluated, therefore, using real-life fashion product updates is not suitable for this model at all times. Therefore, while creating the user preference recommendation system, the user preferences in the existing dataset were evaluated. The model had some problems in classifying categories with similar features. To overcome these obstacles, future research should focus on improving feature extraction techniques or using more complex models.

Limitation of Features: Currently, only visual data is used, which does not allow the model to evaluate additional features such as user comments and ratings. This feature is considered as an additional feature and can be developed in future work.

Data Scope and Variety: The images that the project model suggests are based solely on the dataset used in training. This limits the variety of products that the model can suggest to the dataset. Any fashion item other than the products included in the dataset cannot be recognized by the model and therefore cannot be suggested. This is a limitation that the model may encounter in real-world use, because a product that attracts the user's attention in the real world may not be included in the dataset. However, expanding this in this way would be going beyond the scope of the project. Because the project's goal was to progress with these arrangements within the existing dataset, and this was achieved. Therefore, if it is desired to increase interactivity in different areas in future studies, this feature can also be added.

6.2.4 Future Work

Future research in this area aims to improve classification accuracy and reduce category overlap. These accuracy improvements can be achieved through strategies such as attention processes or transfer learning. Expanding and broadening the datasets might improve the model's capacity to generalize across various fashion categories. Working together with fashion merchants to get actual data may yield insightful information. Future work could incorporate algorithms that strike a balance between novelty and relevance in order to address the problem of suggestion diversity. Methods like hybrid recommendation systems or diversification tactics could improve the efficacy of the system. It will be helpful to carry out further user research to acquire a deeper understanding of user preferences and experiences with the recommendation system. The system's algorithms may be improved and user satisfaction may rise as a result of these investigations.

In addition to these, I aimed to enrich my dataset with user reviews and product descriptions. This is an important step for the model to learn deeper and more meaningful features. However, I would like to point out that these techniques are not implemented in my current code and these features are not among the main goals of the project. This was planned as an additional feature and was not included in the scope of my project. My code currently has a model that only processes basic image and label information. The model uses a convolutional neural network (CNN) designed to classify images and make recommendations. My code works only with image data and labels during the data processing and model training stages. However, the necessary infrastructure has been prepared for adding text data such as user reviews and product descriptions, and this can be improved as stated in the future work section of my report. In the future, the project will process user reviews and product descriptions as text data and include this data in the model: Tokenization can be used to convert text data into numerical data by breaking it down into words or phrases.Vectorization can be used to convert text into vectors and enable the model to learn meaningful information from this data. Word Embeddings can be used to represent words as numerical vectors, thus capturing deeper meanings of texts.

6.3 Conclusion

To sum up, this study's fashion product categorization and recommendation system has achieved its main goals, showing notable progress in terms of classification accuracy and recommendation relevancy. The project's outcomes offer a strong basis for upcoming developments in fashion e-commerce technology. Further improvements to the categorization and recommendation components can be achieved by resolving highlighted limits and investigating potential future work, which will ultimately result in a more reliable and user-centered system.

References

Berg, A., Boehnke, J., Borth, D., & Ulges, A. (2019). Fashion and apparel classification using convolutional neural networks. Journal of Fashion Technology & Textile Engineering, 7(2), 45-59. https://doi.org/10.4172/2329-9568.1000121

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 770-778). https://doi.org/10.1109/CVPR.2016.90

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing, 7(1), 76-80. https://doi.org/10.1109/MIC.2003.1167344

McAuley, J., Targett, C., Shi, Q., & van den Hengel, A. (2015). Image-based recommendations on styles and substitutes. In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 43-52). https://doi.org/10.1145/2766462.2767755

Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), The adaptive web: Methods and strategies of web personalization (pp. 291-324). Springer. https://doi.org/10.1007/978-3-540-72079-9_9

Smith, K., & Dong, C. (2020). Addressing overfitting in fashion classification with data augmentation techniques. Pattern Recognition Letters, 138, 18-24. https://doi.org/10.1016/j.patrec.2020.08.019

Van Meteren, R., & Van Someren, M. (2000). Using content-based filtering for recommendation. In Proceedings of the ECML/PKDD 2000 Workshop on Machine Learning in User Modelling. ECML-PKDD.

Wu, Y., Ahmed, A., Smola, A. J., & Jing, X. (2016). Recurrent neural networks for collaborative filtering. In Proceedings of the 10th ACM Conference on Recommender Systems (pp. 495-498). https://doi.org/10.1145/2959100.2959175

Zhou, Y., Wilkinson, D., Schreiber, R., & Pan, R. (2008). Large-scale parallel collaborative filtering for the Netflix Prize. In Proceedings of the 4th International Conference on Algorithmic Aspects in Information and Management (pp. 337-348). Springer. https://doi.org/10.1007/978-3-540-68880-8_32

Zou, B., & Xu, W. (2019). An improved deep learning approach for fashion image classification. Journal of Machine Learning Research, 20(1), 1-25.