Enabling Circular and Sustainable Clothing in your Wardrobe

Rekha Agali Virupaksha

Universität Koblenz, Koblenz, Germany
rekha_av@uni-koblenz.de

Abstract. In response to the urgent environmental challenges posed by the fast fashion industry, this research introduces a recommendation model aimed at promoting sustainable fashion consumption. Recognizing the inconsistency between consumer environmental consciousness and their fashion choices, the study focuses on incorporating key sustainability attributes—such as the use of eco-friendly materials, ethical production practices, and the longevity of garments—into a recommendation tool that respects individual style preferences. To achieve this, two models were developed : a Random Forest Classifier Model to assess the sustainability of fashion items based on eco-friendly materials, a Content-Based Recommendation model tailored to match individual style preferences with sustainable options.

Keywords: Circular Clothing · Sustainable Fashion · Recommendation Model · Fashion Industry

1 Introduction

In today's fast-paced fashion industry, we are facing a critical challenge: how to make a profit while still protecting the environment. The recent rise of fast fashion has made this challenge even tougher, leading to more waste and higher resource consumption. The rapid evolution of fast fashion has led to an unsustainable cycle of consumption and waste, and it raises serious challenges to environmental sustainability.

Research indicates that the fashion industry is responsible for approximately 10% of global carbon emissions, positioning it as a substantial contributor to climate change. It stands as the second-largest water consumer, while also generating considerable waste, energy consumption, and pollution through microplastics and harmful chemicals. Studies suggest a growing awareness among consumers of the environmental impacts of their fashion choices, with an increasing willingness to adopt more sustainable consumption practices. However, despite this growing consciousness, a significant gap remains between intention and action. Consumers often find it challenging to locate environmentally friendly products, partly due to the overwhelming array of environmental labels and the skepticism surrounding corporate sustainability claims. The lack of a standardized environmental labeling system in the fashion industry exacerbates this issue, making it difficult for consumers to identify the true environmental effects of their purchases [17].

This issue has sparked a strong interest in sustainable fashion, aimed at reducing the environmental impact of making and wearing clothes. Alongside this, the idea of circular fashion, which focuses on creating clothes that can be recycled, reused, or upcycled, supports this eco-friendly approach.

1.1 Motivation

This study is driven by the requirement to shift consumer behavior in the fashion industry towards sustainability. The continuous chase for fast fashion leads to environmental harm due to resource overuse, a higher carbon footprint, and growing amounts of textile waste. With this in mind, this research focuses on creating a technological solution — a recommendation model. This model aims to promote sustainable fashion choices among consumers while ensuring their individual style and preferences remain unaffected.

1.2 Research Questions

To guide this investigation, I aim to explore two critical research questions:

1. What are the key sustainable fashion attributes that should be factored into the recommendation model? This question seeks to identify the essential features that define sustainability in fashion, such as the use of eco-friendly materials, ethical manufacturing processes, and the longevity of garments. Understanding these attributes is crucial for developing a model that accurately assesses the sustainability of fashion items.
2. How can recommendation models be tailored to prioritize sustainable fashion choices without compromising on user preferences? This question addresses the challenge of integrating sustainability into the fabric of consumer choices. It explores the potential of advanced models to balance environmental considerations with individual style and functional needs, thereby facilitating a seamless blend of sustainability and personal preference in fashion recommendations.

1.3 Challenges

The main challenges on developing a recommendation system that prioritizes sustainability while serving to diverse consumer preferences could be described as follows:

- Consumer Awareness and Preference: Balancing consumer desire for trendy, affordable clothing with the necessity for sustainable choices requires refined understanding and strategic guidance.
- Data Complexity: Accurately representing and processing the number of factors that define sustainability in fashion—material origins, manufacturing processes, lifecycle assessments, etc.—demands advanced data handling and algorithmic approaches.

- Adoption and Impact: Ensuring the recommendation system is accessible, user-friendly, and effective in influencing consumer behavior towards making more sustainable choices.

1.4 Contributions

This study has significant contributions towards the sustainable fashion industry. Along with the development of a sustainable fashion recommendation model, this study will also be able to advance the sustainable fashion practices while also providing some practical insights. They are described more aptly in the points below:

- Development of a Sustainable Fashion Recommendation Model: By identifying key sustainability attributes and integrating them into an advanced recommendation framework, this study provides a tool for aligning fashion choices with environmental goals.
- Advancement of Sustainable Fashion Practices: Through a comprehensive analysis of sustainability in the fashion industry, the research highlights the role of technology in promoting eco-friendly consumer behavior.
- Practical Assessment: The application and assessment of the recommendation system offer insights into its effectiveness in encouraging sustainable choices, thereby contributing valuable knowledge to the field of sustainable fashion technology.

2 Related Work

The combination of sustainability and technology in the fashion world has been getting a lot of attention lately. Researchers are looking into how technology can help reduce the harm fashion does to the environment. This section discusses what has been found in the fields of sustainable fashion, circular fashion, and using technology to recommend eco-friendly fashion choices. This background helped shape the tools and methods used in this study.

2.1 Sustainable and Circular Fashion

There are studies on sustainable fashion that focus on making fashion less damaging to the planet. Research by [1] shows different ways the clothing industry is trying to be more eco-friendly, such as, by using materials that are better for the environment and recycling clothes.

Another study by [6] discusses how reusing materials (upcycling) can help reduce waste. These studies stress the importance of understanding the entire process of the fashion industry, from the preparation, production to the consumption of clothes, while creating technological solutions that address these issues.

2.2 Fashion Recommendation Systems

There has recently been an uptrend in fashion retail towards making recommendations more personalized and focused on being eco-friendly. [15] developed an algorithm that makes clothing recommendations based on consumer's preferences and how eco-friendly the choices are.

A review by [8] on fashion recommendation systems shows how combining personal tastes with eco-friendly options can help customers make greener choices. General research on e-commerce recommendation systems by [2], as well as discussions on eco-focused recommendation tools by [10], have helped understand how to build systems that meet customer needs while also caring for the planet.

[17] work stands out as an important contribution to the study of sustainable fashion and technology. Their research looks at how digital tools and online recommendation systems can be better used to promote eco-friendly fashion options to the consumers. They explore ways to make these systems focus more on sustainable products and help users understand the environmental impact of their clothing decisions.

3 Approach

The core objective of this research is to develop a recommendation system that not only aligns with consumers' aesthetic and functional preferences but also promotes the sustainability and circularity of clothing. This section outlines the approach taken to achieve this goal, including data preparation, model development, and the integration of sustainability criteria into the recommendation process.

3.1 Data Preparation

The dataset contains a wide variety of garment types, made from various fibers including recycled materials. To process this dataset, a series of data analysis, cleaning and transformation steps were performed. This included converting string representations of numerical values into floats and encoding categorical variables to make them suitable for machine learning algorithms.

3.2 Analysis and Visualization

To validate the approach and gain insights into the dataset, various analyses were conducted, including material composition analysis and sustainability scoring.

Figure 1 shows the number of products in the dataset with each class label. 'EI' column (Environmental Impact) is considered the class label and it has values from 1 to 5. 1 being low environmental impact and 5 being high environmental impact.

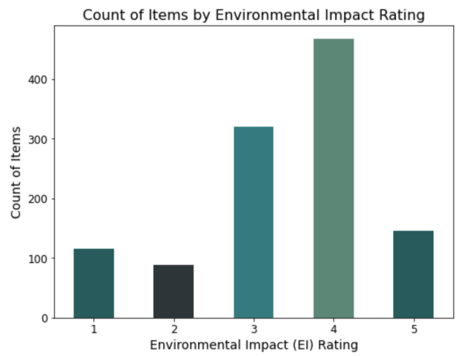

Fig. 1. Number of products in the dataset with each class label.

Figure 2 illustrates the ratio of eco-friendly to non-eco-friendly fabrics in the dataset. Eco-friendly materials, which include natural materials and recycled content, are shown in contrast to non-eco-friendly materials, predominantly synthetic fabrics. This visualization highlights the proportion of materials considered more sustainable and environmentally friendly versus those that are less so, providing a visual summary of the dataset's composition in terms of material sustainability.

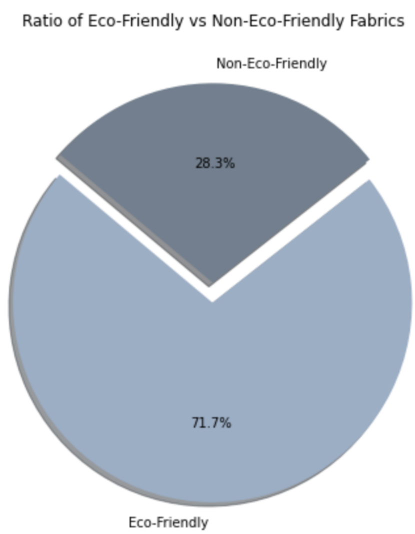

Fig. 2. Ratio of eco-friendly to non-eco-friendly fabrics in the dataset.

Figure 3 displays the total content of each fabric across all items in the dataset. This visualization provides insight into the generality of various materials, highlighting which fabrics are most commonly used in the dataset's clothing items.

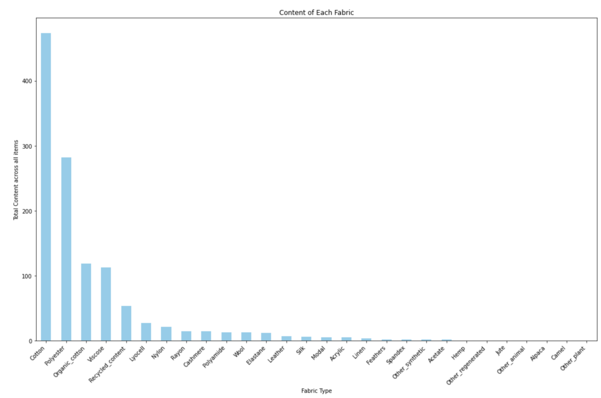

Fig. 3. Total content of each fabric across all items in the dataset.

3.3 Model Development

The recommendation system is developed using a two approaches: a Random Forest Classifier Model to predict item compatibility with user preferences and a Content-Based Recommendation Model to identify items similar to those preferred by the user.

- Random Forest Classifier: This model was trained to classify items based on their likelihood of aligning with a user's preferences, using features such as material composition, brand, and style. The Random Forest Classifier was chosen for its ability to handle a diverse set of features and its robustness in dealing with unbalanced datasets.
- Content-Based Recommendation Model: To complement the classifier and enhance the recommendation process, a content based model was employed. This model calculates the similarity between items using cosine similarity based on their attributes, enabling the recommendation of items that are not only similar to those the user likes but also score high on sustainability measures.

4 Experiments

The research's experimentation stage was important in developing and fine-tuning the recommendation system. This section explores the methods used, from preparing the data to assessing the model, and mentions why certain algorithms and techniques were chosen.

4.1 Data Preprocessing

The first step in our experimentation involved extensive data preprocessing to ensure the dataset was suitable for model training and analysis. This process included:

- Converting Numerical Values: Many material composition percentages were initially represented as strings with commas (e.g., "1,5" for 1.5). A custom function was applied to convert these strings into float values, standardizing the numerical data across the dataset.
- Label Encoding of Categorical Variables: Label Encoding was employed to transform textual categories into numerical values, enabling their use in machine learning models. This approach assigns unique integers to categories, offering a straightforward and dimensionality-preserving method ideal for tree-based models, enhancing computational efficiency and model interpretability. This approach was particularly used in the Random Forest Classifier Model.
- Scaling Numerical Values: To address the wide range of magnitudes present in the dataset's numerical features, a MinMaxScaler was utilized to normalize these values to a common scale between 0 and 1. This transformation is critical for models that are sensitive to the scale of input features, as it ensures that all numerical inputs contribute equally to the model's decision-making process, thereby improving the overall performance and stability of the model during training.
- Encoding Categorical Variables: Given the presence of categorical data (e.g., brand names, style descriptions), one-hot encoding was employed to transform these into a machine-readable format. This method was chosen over label encoding to avoid imposing an artificial ordinal relationship among categories.

```python
                        '''
                         DATA PREPROCESSING
                        '''


#Identifying numerical and categorical columns for preprocessing
numerical_cols = pandas_df.select_dtypes(include=['float64', 'int64']).columns
categorical_cols = pandas_df.select_dtypes(include=['object']).columns

print(numerical_cols)
print(categorical_cols)

#Preprocessing: Setting up transformers for numerical and categorical data
numerical_transformer = MinMaxScaler()
categorical_transformer = OneHotEncoder(handle_unknown='ignore')

#Bundling transformers into a ColumnTransformer
preprocessor = ColumnTransformer(
    transformers=[
        ('num', numerical_transformer, numerical_cols),
        ('cat', categorical_transformer, categorical_cols)
    ])

#Applying the preprocessing to the DataFrame
data_preprocessed = preprocessor.fit_transform(pandas_df)

#Viewing the shape of the preprocessed data for confirmation
data_preprocessed.shape
pandas_df.info()
```

Fig. 4. Data Preprocessing.

4.2 Model Selection and Development

The core of this experimentation centered around the development and testing of two distinct but complementary models:

Random Forest Classifier: This model was chosen for its versatility and robustness in handling both numerical and categorical data. The Random Forest Classifier excels in managing complex interactions between features and does not require extensive parameter tuning to achieve high performance. Its ability to provide insight into feature importance was invaluable for evaluating the significance of sustainability attributes in the recommendations.

```
label_encoders = {}
categorical_columns = pandas_df.select_dtypes(include=['object']).columns
for column in categorical_columns:
    le = LabelEncoder()
    pandas_df[column] = le.fit_transform(pandas_df[column])
    label_encoders[column] = le

#splitting the dataset
feature_columns = pandas_df.columns.drop(['ID'])
target_column = 'EI'
X = pandas_df[feature_columns]
y = pandas_df[target_column]
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

```

Command took 0.09 seconds -- by rekha_av@uni-koblenz.de at 29/03/2024, 23:46:03 on My Cluster

Cmd 10

```
#training random forest model
rf_model = RandomForestClassifier(n_estimators=100, random_state=42)
rf_model.fit(X_train, y_train)

```

Out[25]: RandomForestClassifier(random_state=42)

Command took 0.34 seconds -- by rekha_av@uni-koblenz.de at 29/03/2024, 23:46:04 on My Cluster

Cmd 11

```
#evaluation
predictions = rf_model.predict(X_test)
report = classification_report(y_test, predictions)
print(report)
```

Fig. 5. Random Forest Classifier.

- Feature Selection and Dataset Splitting : I designated the dataset's features and target variable, excluding an identifier column to ensure the model's focus on relevant information. The target variable, representing the Environmental Impact (EI) of each product, guided our classification objective. A split allocated 80% of the data for training and 20% for testing, ensuring an even representation of target classes across both subsets.
- Model Training : The choice of a Random Forest Classifier was motivated by its ability to handle a mix of numerical and categorical data, its robustness to overfitting, and its excellent generalization capabilities. I trained the model with 100 trees, balancing complexity with computational efficiency. The training process involved adjusting the model to recognize patterns correlating feature inputs with the environmental impact categories.

– Model Evaluation : Upon training completion, I evaluated the model's performance using the test set. Metrics such as accuracy, precision, recall, and the F1 score provided a comprehensive view of its efficacy in classifying products based on their environmental impact. These metrics collectively indicated the model's ability to accurately identify products with low environmental impact, highlighting its potential utility in guiding consumers towards more sustainable choices.

Content-Based Recommendation Model: To adjust recommendations more closely to individual user preferences, a content-based filtering approach using cosine similarity was implemented. This model computes the similarity between items based on their attributes, focusing on those with high sustainability scores. The choice of cosine similarity was motivated by its effectiveness in identifying items with similar material compositions and sustainability characteristics, even in the absence of explicit user ratings.

```
#Converting the preprocessed data back to a DataFrame for further manipulation
columns_transformed = preprocessor.transformers_[0][1].get_feature_names_out(numerical_cols).tolist() + preprocessor.transformers_[1][1].get_feature_names_out(categorical_cols).
tolist()
data_preprocessed_df = pd.DataFrame(data_preprocessed, columns=columns_transformed)

#Adjust weights for sustainable materials
sustainable_features = ['Organic_cotton', 'Linen', 'Hemp', 'Recycled_content', 'Reused_content']
weight_multiplier = 2  #example weight multiplier

for feature in sustainable_features:
    if feature in data_preprocessed_df.columns:
        data_preprocessed_df[feature] *= weight_multiplier

#Recompute the cosine similarity matrix with adjusted weights
cosine_sim = cosine_similarity(data_preprocessed_df.values)

display(data_preprocessed_df)


#Calculating cosine similarity among items based on the preprocessed data
cosine_sim = cosine_similarity(data_preprocessed)
```

Fig. 6. Cosine Similarity for Content-Based Model.

– Adjusting Weights for Sustainable Materials : Understanding the importance of promoting sustainable materials, I adjusted the weights for features representing eco-friendly options like Organic Cotton, Linen, and Hemp. By doubling the weight of these sustainable features, I aimed to amplify their influence in the recommendation system, prioritizing items with higher content of these materials.
– Cosine Similarity for Recommendations : To recommend similar items, I computed the cosine similarity among all items based on their preprocessed data. This metric helped me identify products with similar material compositions and environmental impacts. This recommendation function then selects items that are most similar to a given item, excluding the item itself to provide new and relevant suggestions to the user.
– Application and Evaluation : I demonstrated the model's utility by recommending items similar to a given product, showcasing how it can suggest alternatives with potentially lower environmental impacts. Furthermore, I evaluated the model's effectiveness using precision, recall and F1 score, focusing on its ability to recommend items deemed sustainable (with a lower environmental impact indicator).

– Material-Specific Recommendations : The model also allows for material-specific recommendations, enabling users to find products that meet specific sustainability criteria, such as a minimum percentage of organic cotton. This feature supports consumers in making informed decisions aligned with their environmental and ethical values.

5 Conclusion

This research advances the sustainable fashion dialogue by developing a recommendation algorithm that matches consumer style preferences with environmental sustainability. The algorithm, grounded in careful data analysis and the integration of sustainability metrics, proves effective in nudging consumer choices towards eco-friendly alternatives.

Firstly, the integration of a Random Forest Classifier Model and a Content-Based Recommendation Model provided a robust framework for evaluating the sustainability of fashion items. This approach directly responds to the first research question by identifying key sustainable fashion attributes, such as eco-friendly materials, ethical manufacturing processes, and garment longevity. These attributes were carefully factored into the algorithm, ensuring that sustainability considerations are central to the recommendation process.

Secondly, addressing how recommendation algorithms can prioritize sustainable fashion choices without compromising user preferences, the study shows the algorithm's ability to balance environmental considerations with individual style and functional needs. This dual-model approach ensures that recommendations do not merely reflect sustainable options but are also aligned with the user's personal taste and preferences.

Through the development and application of this algorithm, the study effectively meets the challenges of promoting sustainable fashion practices among consumers. It advances the field by demonstrating how technology can be utilised to influence consumer behavior towards more sustainable choices, thereby contributing significantly to the reduction of the fashion industry's environmental impact.

References

1. Pragati Sinha, Monica Sharma, Rajeev Agrawal, A Systematic Review and Future Research Agend For Sustainable Fashion in the Apparel Industry. Article (2022)
2. Samit Chakraborty, Md. Saiful Hoque, Naimur Rahman Jeem, Manik Chandra Biswas, Deepayan Bardhan, Edgar Lobaton, Fashion Recommendation Systems, Models and Methods: A Review. Article.
3. Sapna, Ria Chakraborty, Anagha M., Kartikeya Vats, Khyati Baradia, Tanveer Khan, Sandipan Sarkar, Sujoy Roychowdhury, Recommendence and Fashionsence: Online Fashion Advisor for Offline Experience. Article.
4. Yshaghayegh Shirkhani, Hamam Mokayed, Rajkumar Saini, Hum Yan Chai, Study of AI-Driven Fashion Recommender Systems. Article.

5. Maria Th. Kotouza, Sotirios-Filippos Tsarouchis, Alexandros-Charalampos Kyprianidis, Antonios C. Chrysopoulos, Pericles A. Mitkas, Towards Fashion Recommendation: An AI System for Clothing Data Retrieval and Analysis. Article.
6. Reet Aus, Harri Moora, Markus Vihma, Reimo Unt, Marko Kiisa, Sneha Kapur, Designing for Circular Fashion: Integrating Upcycling into Conventional Garment Manufacturing Processes. Article.
7. Yong-Goo Shin, Yoon-Jae Yeo, Min-Cheol Sagong, Seo-Won Ji, Sung-Jea Ko, Deep Fashion Recommendation System with Style Feature Decomposition. Article.
8. Yashar Deldjoo, Fatemeh Nazary, Arnau Ramisa, Julian McAuley, Giovanni Pellegrini, Alejandro Bellogin, Tommaso Di Noia, A Review of Modern Fashion Recommender Systems. Article.
9. George Stalidis, Iphigenia Karaveli, Konstantinos Diamantaras, Marina Delianidi, Konstantinos Christantonis, Dimitrios Tektonidis, Alkiviadis Katsalis, Michail Salampasis, Recommendation Systems for e-Shopping: Review of Techniques for Retail and Sustainable Marketing. Article.
10. Alexander Felfernig, Manfred Wundara, Thi Ngoc Trang Tran, Seda Polat-Erdeniz, Sebastian Lubos, Merfat El Mansi, Damian Garber, Viet-Man Le, Recommender Systems for Sustainability: Overview and Research Issues. Article.
11. Hyunwoo Hwangbo, Yangsok Kim, Session-Based Recommender System for Sustainable Digital Marketing. Article.
12. Hui Li, Shu Zhang, Xia Wang, A Personalization Recommendation Algorithm for E-Commerce. Article (2013)
13. Reshma Chavan, Debajyoti Mukhopadhyay, A Comparative Study of Recommendation Algorithms in E-Commerce Applications. Article (2017)
14. Badrul Sarwar, George Karypis, Joseph A. Konstan, John Riedl, Analysis of Recommendation Algorithms for E-Commerce. Article (2000)
15. Xiaoyue Cui, An Adaptive Recommendation Algorithm of Intelligent Clothing Design Elements Based on Large Database. Article.
16. Barrie Kersbergen, Sebastian Schelter, Learnings from a Retail Recommendation System on Billions of Interactions at bol.com. Article (2021)
17. Chloe Satinet, François Fouss, A Supervised Machine Learning Classification Framework for Clothing Products' Sustainability. Article (2022)
18. Yehuda Koren, Robert Bell, Advances in Collaborative Filtering. Article (2015)
19. Ekkawut Rojsattarat, Nuanwan Soonthornphisaj, Combining Content-Based and Collaborative Recommendation. Article (2003)
20. Chumki Basu, Haym Hirsh, Willian Weston Cohen, Recommendation as Classification: Using Social and Content-Based Information in Recommendation. Article (1998)
21. Zhi-Dan Zhao, Ming-sheng Shang, User-based Collaborative-Filtering Recommendation Algorithms on Hadoop. Article (2010)

[1]

[1] GitHub repository for source code: `https://github.com/avrekha/recommendation-models`.