Towards Fashion Recommendation: An AI System for Clothing Data Retrieval and Analysis

Maria Th. Kotouza$^{(\boxtimes)}$ (ID), Sotirios–Filippos Tsarouchis,
Alexandros-Charalampos Kyprianidis, Antonios C. Chrysopoulos,
and Pericles A. Mitkas

Electrical and Computer Engineering, Aristotle University of Thessaloniki,
54124 Thessaloniki, Greece
maria.kotouza@issel.ee.auth.com

Abstract. Nowadays, the fashion industry is moving towards fast fashion, offering a large selection of garment products in a quicker and cheaper manner. To this end, the fashion designers are required to come up with a wide and diverse amount of fashion products in a short time frame. At the same time, the fashion retailers are oriented towards using technology, in order to design and provide products tailored to their consumers' needs, in sync with the newest fashion trends. In this paper, we propose an artificial intelligence system which operates as a personal assistant to a fashion product designer. The system's architecture and all its components are presented, with emphasis on the data collection and data clustering subsystems. In our use case scenario, datasets of garment products are retrieved from two different sources and are transformed into a specific format by making use of Natural Language Processes. The two datasets are clustered separately using different mixed-type clustering algorithms and comparative results are provided, highlighting the usefulness of the clustering procedure in the clothing product recommendation problem.

Keywords: Clothing data · Web-crawling · Meta-data analysis · Mixed-type clustering

1 Introduction

The fashion clothing industry is moving towards fast fashion, enforcing the retail markets to design products at a quicker pace, while following the fashion trends and their consumer's needs. Thus, artificial intelligence (AI) techniques are introduced to a company's entire supply chain, in order to help the development of innovative methods, solve the problem of balancing supply and demand, increase the customer service quality, aid the designers, and improve overall efficiency [1]. Recently, an increasing number of projects in the fashion industry make use of AI techniques, including projects run by Google and Amazon.

The use of AI techniques was not possible before the adoption of e-commerce sites and information and communications technology (ICT) systems from the traditional

fashion industry, due to data deficiency. Nowadays, the overflowing amount of data deriving from the daily use of e-commerce sites and the data collected by fashion companies enable solutions related to the fashion design process using AI techniques. Popular fashion houses have provided remarkable AI-driven solutions, such as the Hugo Boss AI Capsule Collection[1], in which a new collection is developed entirely by an AI system, as well as the Reimagine Retail[2] from the collaboration of Tommy Hilfiger, IBM and Fashion Institute of Technology, which aims to identify future industry trends and to improve the design process.

This work focuses on the creative part of the fashion industry, the fashion designing process. To this end, an intelligent and semi-autonomous decision support system for fashion designers is proposed. This system can act as a personal assistant, by retrieving, organizing and combining data from many sources, and, finally, suggesting clothing products taking into account the designer's preferences. The system combines natural language processing (NLP) techniques to analyze the information accompanying the clothing images, computer vision algorithms to extract characteristics from the images and enrich their meta-data, and machine learning techniques to analyze the raw data and to train models that can facilitate the decision-making process.

Several research works have been presented in the field of clothing data analysis, most of them involving clothing classification and feature extraction based on images, dataset creation, as well as product recommendation. In the work of [2], the DeepFashion dataset was created consisting of 800,000 images characterized by many features and labels. In the work of [3], a sequence of steps is outlined in order to learn the features of a clothing image, which includes the following: a) image description retrieval, b) feature learning for the top and bottom part of the human body, c) feature extraction using deep learning, d) usage of pose estimation techniques, and e) hierarchical feature representation learning using deep learning. Other related efforts [4–6] present how to train models using image processing and machine learning techniques for feature extraction.

However, little work has been done in analyzing the meta-data accompanying clothing images. In this work, apart from proposing an AI system which involves many subsystems as part of the clothing design process that can be combined together in order to help the designers with the decision-making process, we emphasize on the data collection, meta-data analysis and clustering techniques that can be applied to improve recommendations.

2 Methods

In this section, we present the proposed decision support system for the designers' creative processes. The system is developed in such a way to be able to model the designer's preferences automatically and be user-friendly at the same, in order to be easily handled by individuals without knowledge of the action planning research field. The system is composed of two interconnected components:

---

[1] https://www.hugoboss.com/fashionstories/digitalisation-is-and-remains-a-big-trend-which-has-already-been-embraced-by-hugo-boss/fs-story-1e6xd6hk2kr8e.html.

[2] https://www.ibm.com/blogs/think/2018/01/tommyhilfiger-ai/.

1. **Offline component**: This component performs (a) data collection from internal and external sources, (b) data storage and management to Databases, and (c) data analysis processes that produce the artificial models which provide personalized recommendations to the end-users.
2. **Online component**: This component comprises mainly the user interface (UI). The users, who are usually fashion designers with limited technical experience, are able to easily set their parameters via the graphical UI, visualize their results and provide feedback on the system results.

The overall system architecture is depicted in Fig. 1, whereas the major subsystems/processes are further analyzed in the following subsections.
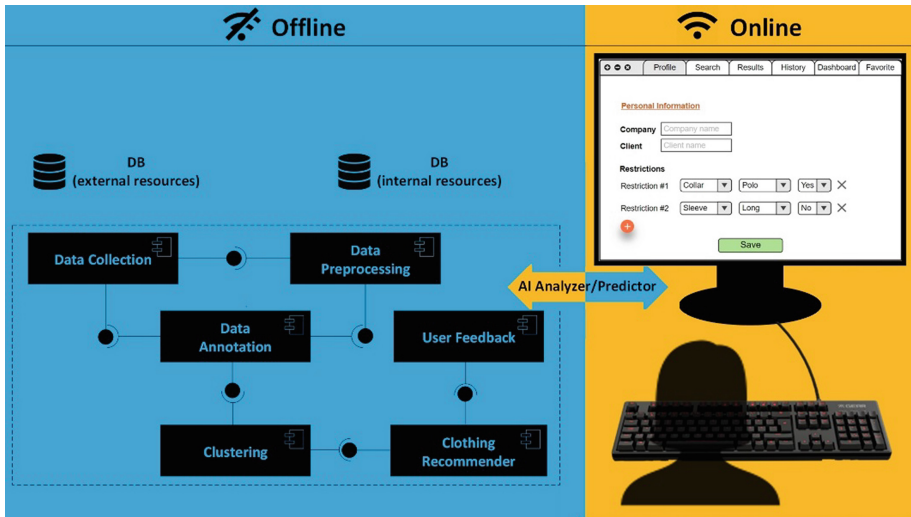

Fig. 1. The proposed system architecture.

2.1 Data Collection

There are two different sources used for training, as well as for the recommendation process: the internal and external data sources.

Internal Data. Each company has its own production line, rules and designing styles that are influenced by the fashion trends. The creativity team usually use an inspiration or starting point based on clothes coming from the company's previous collections and adapt them to the new fashion trends. The internal data are usually organized in relational databases and can be reached by the Data Collection subsystem.

External Data. The most common designers' source for new ideas is browsing on the collections of other popular online stores. To this end, the system includes a web crawler,

the e-shops crawler, which is able to retrieve clothing information, i.e. clothing images accompanied by their meta-data. The online shops that are supported so far are Asos, Shtterstock, Zalando and s.Oliver.

Another important inspiration source for the designers are social media platforms, especially Pinterest and Instagram. To this end, a second web crawler, the social-media crawler, was implemented, which is able to utilize existing APIs and retrieve information from the aforementioned platforms, including clothing images, titles of the post, descriptions and associated tags.

Both crawlers' infrastructure is extendable, so that they can be easily used for other online shops or social media platforms in the future.

2.2 Data Preprocessing

This subsystem is responsible for extracting the clothing attributes from the meta-data accompanying every clothing image. Some of the attributes that are extracted from the available meta-data, accompanied by some valid examples, are presented below:

1. **Product category**: dress, overall, pajamas, shorts, skirts.
2. **Product Subcategory**: jacket, coat, T-shirt, leggings.
3. **Length**: short, long, knee.
4. **Sleeve**: short, ¾ length, sleeveless.
5. **Collar Design**: shirt collar, peter pan, mao collar.
6. **Neck Design**: V-neck, square neck.
7. **Fit**: regular, slim.

For each attribute there is a dictionary, created by experienced fashion designers, that contains all the possible accepted values, including synonyms and abbreviations. NLP techniques are used for word-based preprocessing of all meta-data text. The attributes are extracted using a mapping process between the meta-data and the original attributes. The mapping is achieved by finding the occurrences of the words contained in the dictionaries to the meta-data. In the case of successful matching, the corresponding word is marked as a label to the respective attribute.

2.3 Data Annotation

The Data Annotation process complements the Data Collection and Data Preprocessing modules. It is used to enrich the extracted data with common clothing features that can be derived from images using computer vision techniques. Examples of clothing attributes that can be extracted from images include color, fabric and neck design.

It is widely known that color has the biggest impact on clothing, as it is related to location, occasion, season, and many other factors. Taking into consideration its importance, an intelligent computer vision component was implemented. This component has the capability to distinguish and extract the five most dominant colors of each clothing image. More specifically, the color of a clothing image is represented by the values of the RGB channels and its percentage, the color ranking specified by the percentage

value and the most relevant general color label to the respective RGB value. The rest of the clothing attributes are extracted using deep learning techniques. Each attribute is represented by a single value from a set of predefined labels.

2.4 Clustering Based on Meta-Data

After the Data Collection and Annotation processes, all the data are available in a common format (row data) that can be analyzed using well-known state-of-the-art techniques. A common technique to organize data into groups of similar products is clustering. Clustering can speed up the recommendation process, by making the look-up subprocess quicker when it comes to significant amount of data. A practical example is a case where a user makes a search at the online phase: the system can limit the data used for product recommendation to those that are included in the clusters characterized by labels related to the user's search.

Several clustering algorithms can be used depending on the type of the data. Clothing data can be characterized by both numerical (i.e. product price) and categorical features (i.e. product category) in general. A detailed review of the algorithms used for mixed-type data clustering can be found in [7]. The algorithms can be divided in three major categories: a) partition-based algorithms, which build clusters and update centers based on partition, b) hierarchical clustering algorithms, which create a hierarchical structure that combines (agglomerative algorithms) or divides (division algorithms) the data elements into clusters, based on the elements' similarities, and c) model-based algorithms, which can either use neural network methods or statistical learning methods, choose a detailed model for every cluster and discover the most appropriate model. The algorithms that we use in this paper are as follows:

1. Kmodes[3]: A partition-based algorithm, which aims to partition the objects into k groups such that the distance from objects to the assigned cluster modes is minimized. The distance, i.e. the dissimilarity between two objects, is determined by counting the number of mismatches in all attributes. The number of clusters is set by the user.
2. Pam[4]: A partition-based clustering algorithm, which creates partitions of the data into k clusters around medoids. The similarities between the objects are obtained using the Gower's dissimilarity coefficient [8]. The goal is to find k medoids, i.e. representative objects, which minimize the sum of the dissimilarities of the objects to their closest representative object. The number of clusters is set by the user.
3. HAC[5]: A hierarchical agglomerative clustering algorithm, which is based on the pairwise object similarity matrix calculated using the Gower's dissimilarity coefficient. At the beginning of the process, each individual object forms its own cluster. Then, the clusters are merged iteratively until all the elements belong to one cluster. The clustering results are visualized as a dendrogram. The number of clusters is set by the user.

---

[3] https://www.rdocumentation.org/packages/klaR/versions/0.6-14/topics/kmodes.

[4] https://www.rdocumentation.org/packages/cluster/versions/2.1.0/topics/pam.

[5] https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/hclust.

4. FBHC[6]: A frequency-based hierarchical clustering algorithm [9], which utilizes the frequency of each label that occurs in each product feature to form the clusters. Instead of performing pairwise comparisons between all the elements of the dataset to determine objects' similarities, this algorithm builds a low dimensionality frequency matrix for the root cluster, which is split recursively as one goes down the hierarchy, overcoming limitations regarding memory usage and computational time. The number of clusters can be set by the user or by a branch breaking algorithm. This algorithm would iteratively compare the parent clusters with their children nodes, using evaluation metrics and user-selected thresholds.
5. VarSel[7]: A model-based algorithm, which performs the variable selection and the maximum likelihood estimation of the Latent class model. The variable selection is performed using the Bayesian information criterion. The number of clusters is determined by the model.

2.5 Clothing Recommender and User Feedback

The Clothing Recommender is the most important component of our system, since it combines all the aforementioned analysis results to create models that make personalized predictions and product recommendations. The internal and external data, the user's preferences, and the company's rules are all taken into consideration.

Moving on to the online component, the UI enables the designer to search for products using keywords. The extracted results can then be evaluated by the designer and the preferred products can be saved on their dashboard over time and for each product search. If the user is not satisfied by the recommendations, they have the ability either to renew their preferences or ask for new recommendations.

The offline and the online components are interconnected by a subsystem that is responsible for implementing the models feedback process. The user can approve or disapprove the proposed products based on their preferences, and this information is transmitted as input to a state-of-the-art Deep Reinforcement Learning algorithm, which assesses the end user's choices and re-trains the personalized user model. This is an additional learning mechanism evolving the original models over time, making the new search results more relevant and personalized.

3 Experimental Results

A real-life scenario is provided as a use case, in order to highlight the usefulness of the clustering procedure in the clothing product recommendation. Our team is collaborating with a fashion designer working for the Energiers Greek retail company, who is interested in designing the company's collection for the new season. She uses the garments designed and produced by the company in the previous season as a source of inspiration, combined with the Assos e-shop current collections.

---

[6] https://github.com/mariakotouza/FBHC.

[7] https://www.rdocumentation.org/packages/VarSelLCM/versions/2.0.1/topics/VarSelCluster.

3.1 Datasets

In this direction, the Company dataset was created by extracting the fashion products from the previous season from the company database, and the relevant E-shop dataset was retrieved using a web crawler. A total of 4674 images were collected by the e-shop crawler for the season winter 2020, by making queries involving different labels of the attributes *Product Category, Length, Sleeve, Collar* and *Fit*. The meta-data of the retrieved images and a pointer to the image location were stored in a relational database. The meta-data were tokenized and split into columns, by assigning values in the desired attributes, after preprocessing plain text using NLP techniques.

3.2 Data Clustering

In this section, the experimental results on the Company and E-shop datasets using the Kmodes, Pam, HAC, FBHC and VarSel algorithms are presented. The results are evaluated using four internal evaluation metrics:

a. **Entropy**, which quantifies the expected value of the information contained in the clusters.
b. **Silhouette**, which validates the consistency within the clusters.
c. **Within sum of square error (WSS)**, which is the total distance of data points from their respective cluster centroids and validates the consistency between the objects of each cluster.
d. **Identity** [9], which is expressed as the percentage of data contained in the cluster with an exact alignment regarding the feature's labels.

Lower values of Entropy and WSS, and higher values of Silhouette and Identity indicate better clustering results. The clustering results differ according to the applied clustering algorithms. Table 1 shows the normalized mutual information [10] of the algorithms that were tested, in a pairwise fashion. The values show some variance, with most of them being around 30%. It is worth mentioning that the Pam and FBHC algorithms share information that reaches 59.87%, which is something that can enhance their reliability. On the other hand, the least amount of information is shared between the clusters formed by VarSel and Kmodes, FBHC. The main reason seems to be that VarSel algorithm has automatically identified only 3 clusters, whereas the rest of them have formed 6 clusters. The number of clusters (k) for Kmodes, Pam, HAC and FBHC

Table 1. The normalized mutual information of the algorithms tested on the Company dataset.

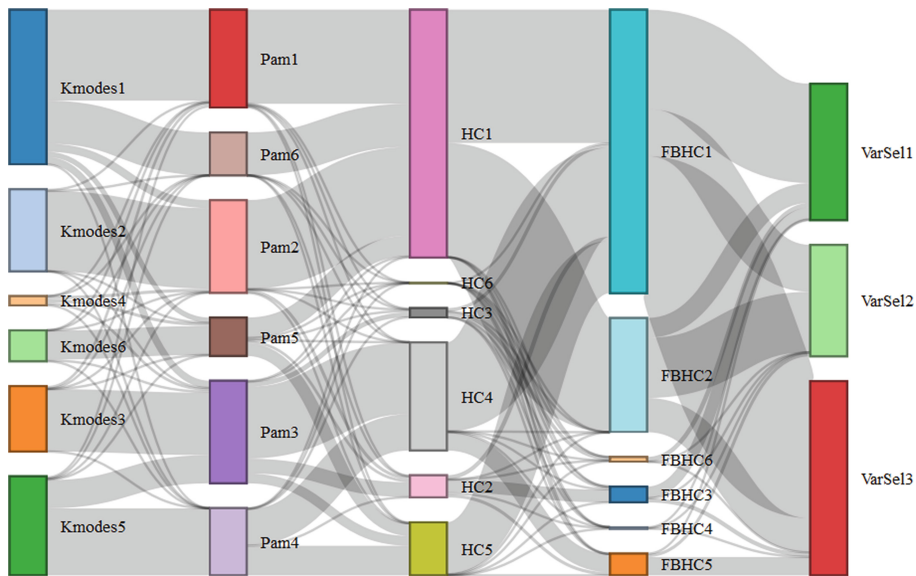|	Kmodes	Pam	HAC	FBHC	VarSel
Kmodes	1		0.3855	0.3036	0.3535	0.0923
Pam	0.3855	1	0.3740	0.5987	0.2835
HAC	0.3036	0.3740	1	0.3082	0.3239
FBHC	0.3535	0.5987	0.3082	1	0.2682
VarSel	0.0923	0.2835	0.3239	0.2682	1

Fig. 2. The Sankey diagram of the clustering results achieved by the algorithms tested on the Company dataset.

was given as input parameter to the algorithms, after experimenting with varying values of k (2 to 12 clusters) and calculating the WSS and Silhouette metrics.

A graphical representation of the information shared across the clusters created by the different algorithms can be seen on the Sankey diagram depicted in Fig. 2. The figure makes clear that the Pam algorithm uniformly distributes the data objects across the 6 clusters, whereas Kmodes clustering results follow a normal distribution. The distributions of those two portioning algorithms seem to be close. The VarSel algorithm normally distributes the objects in a similar fashion, but in this case only 3 clusters are created. On the other hand, the hierarchical algorithms create two large size clusters, where the majority of the objects are assigned to, and four significantly smaller clusters.

Table 2 reports the comparison results of the clustering algorithms based on the values that they achieved at the evaluation metrics. The average values of the evaluation metrics are presented. The best results achieved by an algorithm are highlighted as boldface, whereas the second highest results are presented in italics. The table makes clear that there is not a unique best algorithm that achieves the best results in all the evaluation metrics, so the algorithm's selection depends on the application needs. The hierarchical algorithms achieved better results at the Entropy and Identity metrics, which means that the number of labels characterizing each feature in a cluster is small, whereas the partition-based algorithms outperform at the metrics that concern the distances between the objects of each cluster. Once again it is proved that the Pam algorithm uniformly distributes the data across clusters, and this is the reason why we select this algorithm for the rest of the analysis in this paper. A 2-dimensional representation of the distribution of the data into the six groups obtained by Pam can be seen in Fig. 3.

Table 2. The evaluation results for the Company dataset using various clustering methods.

	Kmodes	Pam	HAC	FBHC	VarSel
# Clusters	6	6	6	6	3
Entropy	0.3279	0.2500	*0.2050*	**0.1889**	0.5159
Silhouette	*0.3479*	**0.4660**	0.2093	0.0753	0.2478
WSS	*0.2800*	**0.2300**	0.4100	0.4900	0.4200
Identity	0.1296	0.1481	*0.1851*	**0.3888**	0.0740

The centroids of the Company dataset extracted by the Pam algorithm are depicted in Table 3 and Table 4 accordingly. The centroids are determined as the most frequent attribute values of the row data for each cluster. A more detailed representation of the groups' consistency for the attributes *Product Category* and *Gender* can be obtained using a heatmap (Fig. 4). By analyzing the consistency of each group and the distribution of the labels across the groups in the two datasets, one can observe that the Company dataset is characterized by six major categories, i.e. Set, Bermuda, Blouse for Men and Women, Dress, and Leggings. On the other hand, the E-shop dataset is characterized by Dress, Shirt, Trousers, Set, Romper, and Cardigan.

As for the rest of attributes, most of the products are characterized by Short Length in the Company dataset, whereas in the E-shop dataset the Medium and Knee Length are more frequent. The tables make clear that the Collar attribute has many missing values, so a good practice will be to recognize this attribute at the Data Annotation subsystem, using computer vision techniques. As for the Fit, the Regular Fit value is the most common in both datasets.

Therefore, when the fashion designer is interested in designing a red dress, she can set the parameters for the product category and the color through the UI of the system and press the search button. The system will then refer to the Company database and filter only the products that are included in the Group 4 created by the offline clustering procedure. The same procedure will be followed to filter only the products that belong to Group 1 in the e-shop's database. The two groups are then combined and the system can select only those products with the label "red" at the Color attribute. Next, this subset can be filtered even more according to the designer's additional preferences and the fashion trends to extract personalized recommendations. Finally, the designer can interact with the system to evaluate (grade) each recommended product, create her dashboard or even ask for new recommendations results if she is not satisfied at all.

4   Discussion

In this work, an intelligent system that automates the typical procedures followed by a fashion designer is described. The system can retrieve data from online sources and the designer's company database, transform plain text accompanying images into clothing features using dictionary mapping and NLP techniques, extract new features from the images using computer vision, and store all the information into a common format in a
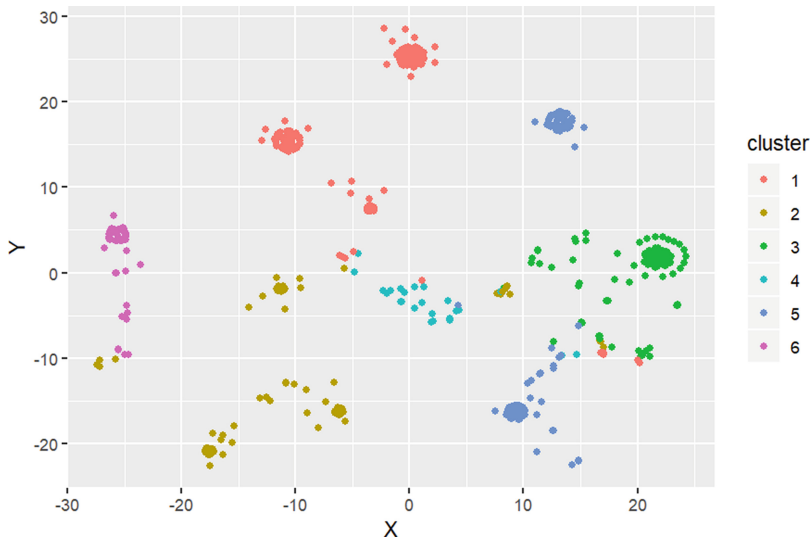
Fig. 3. A 2-dimensional representation of the distribution of the Company dataset into the six groups obtained by Pam.

Table 3. The clustering centroids using the Company dataset using Pam.

	Category	Gender	Length	Sleeve	Collar	Neck	Fit
C1	Set	Man	Short	Long	Shirt	Off shoulder	–
C2	Bermuda	Man	Short	Sleeveless	Flat Knitted Rib	Round	Regular
C3	Blouse	Man	Medium	Short	Shirt	Round	Regular
C4	Dress	Woman	Short	Sleeveless	Flat Knitted Rib	Round	–
C5	Blouse	Woman	Short	Short	–	Round	Regular
C6	Leggings	Woman	Capri	–	–	–	Slim

Table 4. The clustering centroids using the E-shop dataset from the web-crawler using Pam.

	Category	Gender	Length	Sleeve	Collar	Neck	Fit
C1	Dress	Woman	Medium	Raglan	–	Halterneck	Regular
C2	Shirt	Man	Medium	Long	Stand-up	Collar	Regular
C3	Trousers	Man	Medium	Flared	–	Collar	Cargo
C4	Set	Woman	Knee	Flared	–	Collar	Regular
C5	Romper	Woman	Knee	Flared	–	Off shoulder	Regular
C6	Gardigan	Woman	Knee	Long	–	V-neck	Regular

Fig. 4. The distribution of the features' labels across the six different clusters for (a) the Company and (b) the E-shop datasets using Pam.

relational database. The processed data can then be handled by state-of-the-art machine learning techniques including clustering, prediction models, and recommender systems. The paper focuses on presenting the system's architecture, emphasizing on the data collection and transformation processes, as well as the clustering procedures that can be used to organize the row data into groups. A real-life use case scenario was also presented, showing the usefulness of the clustering procedure in the product recommendation problem. Future work involves the augmentation of the Data Annotation process, enabling the extraction of new relevant attributes from non-annotated images. Additionally, the extended use of the products prices and the products' sales history can enrich the model creation process significantly, leading to more reasonable and personalized suggestions for the designers.

Additional steps can be taken in the direction of the improvement of the user-friendliness and the capabilities of the UI, which will be utilized by the designers to enter their preferences, search products, save the system's products recommendation

and create dashboards. Finally, an extended set of experiments using new datasets and methods are needed, whereas testing and evaluation of the recommended products are going to be done by fashion designers in more real-life use case scenarios.

References

1. Guo, Z.X., Wong, W.K., Leung, S.Y.S., Li, M.: Applications of artificial intelligence in the apparel industry: a review. Text. Res. J. **81**(18), 1871–1892 (2011)
2. Liu, Z., Luo, P., Qiu, S., Wang, X., Tang, X.: DeepFashion: powering robust clothes recognition and retrieval with rich annotations. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), no. 1, pp. 1096–1104 (2016)
3. Li, R., Feng, F., Ahmad, I., Wang, X.: Retrieving real world clothing images via multi-weight deep convolutional neural networks. Cluster Comput. **22**(3), 7123–7134 (2017). https://doi.org/10.1007/s10586-017-1052-8
4. Hidayati, S.C., You, C.W., Cheng, W.H., Hua, K.L.: Learning and recognition of clothing genres from full-body images. IEEE Trans. Cybern. **48**(5), 1647–1659 (2018)
5. Kalantidis, Y., Kennedy, L., Li, L.-J.: Getting the look: clothing recognition and segmentation for automatic product suggestions in everyday photos. In: Proceedings of the 3rd ACM Conference on International Conference on Multimedia Retrieval, pp. 105–112 (2013)
6. Vittayakorn, S., Yamaguchi, K., Berg, A.C., Berg, T.L.: Runway to realway: visual analysis of fashion. In: Proceedings - 2015 IEEE Winter Conference on Applications of Computer Vision, WACV 2015, pp. 951–958 (2015)
7. Balaji, K., Lavanya, K.: Clustering algorithms for mixed datasets: a review. Int. J. Pure Appl. Math. **18**(7), 547–556 (2018)
8. Gower, J.C.: A general coefficient of similarity and some of its properties. Biometrics **27**(4), 857–871 (1971)
9. Kotouza, M.Th., Psomopoulos, F.E., Mitkas, P.A.: A dockerized framework for hierarchical frequency-based document clustering on cloud computing infrastructures. J. Cloud Comput. **9**(1), 1–17 (2020)
10. Strehl, A., Ghosh, J.: Cluster ensembles—a knowledge reuse framework for combining multiple partitions. J. Mach. Learn. Res. **3**, 583–617 (2002)