Article

Learning Context-Aware Outfit Recommendation

Ahed Abugabah [1], Xiaochun Cheng [2],* and Jianfeng Wang [3],*

[1]   College of Technological Innovation, Zayed University, Abu Dhabi 144534, UAE; Ahed.Abugabah@zu.ac.ae
[2]   Department of Computer Science, Middlesex University, London NW44BE, UK
[3]   School of Data and Computer Science, Sun Yat-Sen University, Guangzhou 510006, China
*   Correspondence: x.cheng@mdx.ac.uk (X.C.); wangjf7@mail2.sysu.edu.cn (J.W.)

check for updates

Abstract: With the rapid development and increasing popularity of online shopping for fashion products, fashion recommendation plays an important role in daily online shopping scenes. Fashion is not only a commodity that is bought and sold but is also a visual language of sign, a nonverbal communication medium that exists between the wearers and viewers in a community. The key to fashion recommendation is to capture the semantics behind customers' fit feedback as well as fashion visual style. Existing methods have been developed with the item similarity demonstrated by user interactions like ratings and purchases. By identifying user interests, it is efficient to deliver marketing messages to the right customers. Since the style of clothing contains rich visual information such as color and shape, and the shape has symmetrical structure and asymmetrical structure, and users with different backgrounds have different feelings on clothes, therefore affecting their way of dress. In this paper, we propose a new method to model user preference jointly with user review information and image region-level features to make more accurate recommendations. Specifically, the proposed method is based on scene images to learn the compatibility from fashion or interior design images. Extensive experiments have been conducted on several large-scale real-world datasets consisting of millions of users/items and hundreds of millions of interactions. Extensive experiments indicate that the proposed method effectively improves the performance of items prediction as well as of outfits matching.

Keywords: visual style; context-aware; preference analysis; fashion recommendation

1. Introduction

As we log into a shopping website, just like billions of other users, we face a large and diverse range of product information. Intelligent apparel recommendation systems have been widely deployed in Web applications that intend to assist customer search, select and match clothes to support various services online. Taobao is the largest online consumer-to-consumer platform in China. Users come to Taobao to select and combine items of clothes to form a coordinated outfit. If we want to buy a coat or pair of trousers, we need to pick from thousands of items, and it takes several hours to get the right clothes. Without a doubt, fashion recommendation systems have become a key component in many online shopping applications, spanning from visual information to style matching recommendation.

Fashion products are different from general consumer goods, they have a complex pattern design structure, unique style and diverse color. Because of large deformations, occlusions, and discrepancies of clothes, fashion outfits are closely related to material information, texture information and detail information. Thus, consequently, it is important to analyze the complex relationships between items and user preferences such as user style, the perceived image which equally impacts on the consumer.

Recently, the Internet and widespread use of fashion shopping platforms and fashion communities such as Taobao, Lookbook (https://lookbook.nu/) and Chictopia (http://chictopia.com/), has brought

increasing attention to fashion recommendation. In Taobao, the largest online consumer-to-consumer platform in China, a new application iFashion has been created to support fashion outfit recommendation. Approximately 1.5 million content creators were actively supporting Taobao as of 31 March 2018 (https://www.alibabagroup.com/en/news/presspdf/p180504.pdf). Fashion outfit recommendation has become more and more important for modern consumers and has thus attracted interest from the online retail industry. Hu et al. [1] maps feature vectors from feature space to some low-dimensional latent space. This gradient-based method learns nonlinear functions and successfully uses the visual features of fashion to solve the problem of automatically recommending clothing to users based on personal preferences. Veit et al. [2] proposed a self-learning style feature transformation method, which can solve the problem of matching different types of objects in the style space from the image of goods to the learning of style space and the feature transformation from the image of goods to the latent space.

Previous works [2–4] have transformed visual information into embedding vectors that enhance customer service and improve sales for fashion retailers. Though these methods have improved the recommendation performance that recommend new products for potential consumers according to the customer body, occasion needs and fashion coordination. In fact, for some types of items, apparel has dynamic deformations caused by material and movements, and is hard to capture its full visual features. Moreover, their visual appearances can greatly bias the user's preference towards them. Therefore, modeling user preference with interacted images still suffers from some inherent issues: (1) Due to the differentiation of individual backgrounds and fashion involvements, the perception and judgment of apparel meanings may not result in consistency, so it is hard to measure and effectively learn in a machine learning model. (2) Most methods assume that the user's interest is constant and generates a static interpretation. In fact, user preferences are always changing, and they have a unique feeling when they are in different environments and states. Facing tens of thousands of clothes styles in a database in current online stores, it is challenging not only recommending similar products to meet users' current style, but also finding appropriate clothes to match individual needs. In this paper, we take the personalized characteristics of users and the relationship of items into account.

For the above problems, we integrate the historical interaction between users and items with the interrelationships among items in multiple items, and through information exchange, users have a better understanding of fashion recommendation preferences. We propose a fashion learning model based on the dynamic change of user preferences and visual compatibility relationships, as shown in Figure 1. The proposed model explains three tasks. Task 1: predicting the items of user interest. Task 2: generating an outfit based on user text/image. Task 3: recommending fashion items that match different categories of clothes.
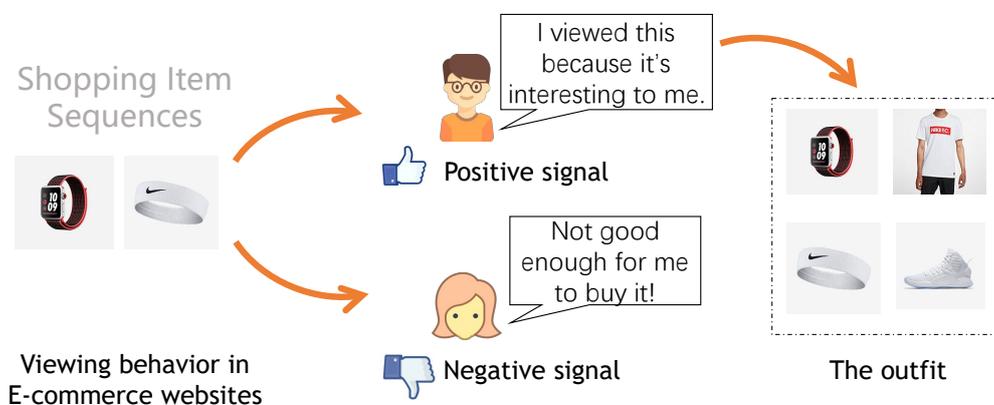

Figure 1. Three tasks of fashion items matching.

By analyzing the historical interactions between users and items, and visual attributes [5], we propose a model that incorporates visual compatibility relationships as well as historical interactions

between users and items, aiming to integrate multiple item relations for better recommendation. We use neural networks to extract image features, and model compatibility through the relationship between product image features. In the model, several additional layers of understanding of the visual elements of product styles are added. To evaluate the model, large data sets [6] from Amazon (http://jmcauley.ucsd.edu/data/amazon/) and Epinions.com (www.Epinions.com)were used to test the collocation between apparel items. The results show that the proposed method capture user interests, and match the visual style of clothing consistently.

To integrate the interest items and the user's historical interaction into the model learning procedure and help to predict their next action based on the context of what they have done recently, we propose a fashion recommendation method which explores the influence of different visual items between the items compatible with external outfits and the items in the target field. In addition, the relationship is effectively applied to the recommendation system, which improves the user experience on the shopping website [7], arouses the second purchase desire of users, and promotes the sales of clothing. The proposed model not only improves the compatibility between items but also researches higher-level attributes relationships between items. The contributions of our work are three-fold as follows:

- Incorporating visual style information into historical interactions between users and items, we explore how information interchange on various items reflects user preferences, so as to capture user preference on visual information.
- Inspired by visual style compatibility, we contribute a simple and effective outfits solution that integrates a personalized higher-order Markov modeling into a visual-based recommendation framework, the proposed method is context-aware to capture the importance of compatibility information.
- The user's dynamic interest is fused with his or her fashion style into a unified personalized recommendation model, which can provide the users with visual style compatibility.

The content of this paper is designed as follows. In Section 2, we briefly explain some related research. Section 3 introduces our method in detail and explains notations. The experiments and methods for analyzing the results are demonstrated in Section 4. Lastly, conclusions and future work are outlined in Section 5.

2. Related Work

2.1. Visual Features

Fashion outfit is formed by people's random collocation of clothes and is strongly subjective [8], which makes it difficult to determine with a calculation model. If the style is identified by low-dimensional visual similarity [2,9], it may be too specific to detect similar-style images, and the hand-made style may be too abstract to capture subtle style differences. The clothing style is different from clothing shape (sweater, dress) or clothing attribute [10–12]. In the study of clothing style recognition, the early method [13,14] explored supervised learning and classify the style according to the user's identity information [15].

Some work has been done on garment visual feature modeling [16], including garment matching [4,17] and visual recommendation [16]. McAuley et al. [6] proposed personalized matching based on visual features. He et al. [18] combined the characteristics of visualization and proposed a framework of Bayesian Personalized Ranking (BPR), which improved the performance of product recommendation in implicit feedback scenarios [19]. On this basis [6], He et al. [18] proposed the VBPR method, and further expanded the dynamic dimension to simulate the visual evolution of the fashion trend of visual recommendation.

2.2. Sequence Recommendation

In many current recommendation systems, sequence recommendation plays a critical role [18,20,21]. Sequential recommendation methods usually discover the internal connections between previously purchased items to accurately predict the next item. Generally, sequence models are divided into two main types: (1) the order-based models view user sequences as item orders, mainly mining the relationship between commodity sequences, such as Markov Chains [22], recurrent neural network RNN [23,24], convolutional neural network [CNN] [25,26] and attention mechanism [27]; GRU4Rec used a gated recurrent unit (GRU) to build session-based recommendations [28] Modular click sequence, and the improved version further improves its Top-N recommendation performance [29]. However, in each time step, the RNN takes the state and current operation of the last step as its input. These dependencies reduce the efficiency of the RNN [28]. (2) Based on the sequence perception model, it is only a model for establishing interaction records between users and commodities, without considering the visual characteristics of the image. For example, Rendle et al. [21] proposed a model based on user–item interaction with implicit feedback pairing learning algorithm. Jing et al. [30] proposed a probabilistic matrix decomposition model based on product recommendation, which recommends Flickr photos to evaluate products for users.

Overall, limited work has been done to explore the effects of visual style to the different items. Unlike identifying a piece of clothing or its attributes (blue, flower scent), style required a more advanced concept of how people choose clothing, which is a trend. Most the previous studies were based on the extraction of visual features [2] of clothing products, i.e., by determining the complementary or substitutional relationship of clothing from image features. McAuley [18] determined the relationship between the appearance of a product by capturing the largest possible dataset. The modeling of such a cognitive concept often used the method of labeling images manually, using small data sets and more complex procedures to avoid the problem of over fitting.

3. User Interest Analysis Model

In this section, our model is mainly composed of three components—the user's preferences built on user's historical feedback, user style analysis, and scene-based outfits. By combining the impact of shifting user preferences and visual compatibility relationship between items, we tap the user's preferences from the user's recent browsing, clicks, purchases and other implicit feedback. Specifically, user set and item set are denoted by U and I, such that the collection I_u^+ denotes items that interact with the user u, and $I_u^+ = I - I_u^-$ represents unobserved interactive items.

3.1. Problem Formulation and Notation

In recommendation system, the common methods focus on features of each item, reflecting their content similarity, but seldom consider the higher-level attribute relationships between items. By determining the relationship [31] between items, we can make full use of the user's behavior data to make a reasonable fashion recommendation for users [32]. The dynamic preference model is based on observing and analyzing user behavior, to more accurately reflect the changes in user preferences. To provide the user with personalized outfits, we propose a fashion recommendation model based on the user's dynamic preferences, visual style and attributes of the items. Here we represent a set of users as $\mathcal{U}$, where $\mathcal{U} = (u_1, u_2, ..., u_m)$. We denote a set of items as $\mathcal{I}$, where $\mathcal{I} = (i_1, i_2, ..., i_n)$. $\vec{p}_i$ represents the item i related latent vector, $f_i \in \mathbb{R}^K$, and K represents the dimension of the vector for each user/item. η represents the user's personalized weighting vector, where $\eta = (\eta_1, \eta_2, ..., \eta_u)$. The weighting vector was represented as α, $\alpha \in \mathbb{R}$.

3.2. Preferences-Based Sequence Model

The objective of our task is to predict the sequential behavior of users given the visual information [33] to recommend the top-N items for the user u, user u's representation includes

two part: one is long-term preferences modeling through historical data mining; the other is short-term preferences model that reflects in the data generated by recent browsing and interactions.

Given a user u and the last item they interacted with $i \in I$. The relationship between user u and item i, $\hat{p}_{u,i}$ is estimated to predict the probability that u transitions to another item i by the inner product of x_u and y_i. For convenience, we use the conventional inner product as shown in Equation (1).

$$\hat{p}_{u,i} = < x_u, y_i >, \tag{1}$$

where the inner product models the 'affinity' between latent user factors and latent item factors $x_u \in R^{k1}$ and latent item factors $y_i \in R^{k2}$. The users and items are projected into a low-dimensional space that reflects the extent to which user u prefers image i. The preference of a user u toward an item i is predicted by Equation (2):

$$\hat{p}_{u,i} = < m_i, m_j > + < x_u, y_i >, \tag{2}$$

where $< m_i, m_j >$ models the 'continuity' between item i and the previous item j. $m_i, m_j \in r^{k2}$.

In Equation (2), m_i and m_j are latent vector representations of commodity i and j respectively. The historical behaviors of users reflect personal preferences; for each product, we extract features from the products with interactive behaviors, and the aggregation is propagated forward to a unified user representation vector. For commodities that do not engage in interactive behaviors, they are predicted by calculating their similarity to commodities with interactive behaviors. $< m_i, m_j >$ is calculated to predict the relationships of item i and the previous item j.

$$\hat{p}_{u,i} = (\eta + \eta_u) < m_i, m_j > + < x_u, y_i > . \tag{3}$$

where the first represents the drift of the user's short-term preferences. It can be divided into two levels of relationships. First, attention is used to distinguish which types of relationships are more important, and then attention is used to determine which types of historical behavior have more important relationships.

The relative weight of dynamic preference is controlled by scalar η_u. The global bias η controls the dynamic changes of the user's preferences. m_j, n_j represent the interaction between user u and item j.

$$\hat{p}_{u,i} = \sum_{j \in x_u} < x_u, y_i > + (\eta + \eta_u) < m_i, m_j > . \tag{4}$$

Based on Equation (4), we add a regularized dynamic factor η_i, where η_i is $\in \mathbb{R}$. We get a new formula Equation (5).

$$\hat{p}_{u,i} = \beta_j + \frac{1}{|i_u^+|^\alpha} \sum_{j \in x_u} < x_u, y_i > + (\eta + \eta_u) < m_i, m_j > \tag{5}$$

In the time range, the interaction between the user and the product is tracked by Equation (6), and the recommendation list is formed by analyzing the user's shopping record and ranking the product that matches the user's preferences and style.

3.3. Fashion Recommendation with Visual Compatibility Relationship

As mentioned earlier, visual factors play a very important role in the fashion field and affect user behavior. To enhance the feature representation of commodities and better predict user preferences. We extract the visual features of commodities through a preprocessing program and embed them into the visual feature space, and further enhance the user's expressive ability by scoring all the previous images. Inspired by [18], our method adopts the efficientNet-B7 to extract visual features f_i from the raw image.

$$h_i = w_h f_i + b, \tag{6}$$

where h_i is termed as image embedding, and w_h and b are parameters to be learned. Based on the neural network incorporating visual perception to analyze the user's dynamic preferences for different commodities. In the training, the visual information of the commodities is combined to embed the items representation, forming a unified representation for the user–item pair.

The result is fused to jointly account for the two representations. The motivation based on neural networks for incorporating the user-aware coefficient is to capture users' dynamic preferences over different components. We compose a representation for the user–item pair by combining the item embedding and their visual features. Compared to BPR-MF, there are now two sets of parameters to be updated: (a) the non-visual parameters, and (b) the newly introduced visual parameters. Non-visual parameters can be updated in the same form as BPR-MF (therefore are suppressed for brevity), while visual parameters are updated according to:

$$\hat{p}_{u,i} = \beta_j + \frac{1}{|i_u^+|^\alpha} \sum_{j \in x_u} <x_u, y_i> + (\eta + \eta_u) <m_i, m_j> \\ + \gamma_\mu^T (h_i), \tag{7}$$

where γ_μ^T represents a latent visual preferences for user u, and f_i is projected into a *K*-dimensional latent space, which the style space explains in user preference.

In the above objective function, the first term is used to maximize the possibility of the user's implicit feedback and predict the user's preference. The middle item is used to calculate the compatibility relationship between commodity i and commodity j, and the last term is to regularize the parameters to avoid over fitting.

For each user u, the maximum posterior probability (MAP) is inferred by optimizing the model parameters θ. γ is the visual style of the product. $p(\theta, i)$ indicates the probability that the item j matches the clothing item i. Considering the compatibility of the visual features, we select random samples from all training examples to extract the related parameters. To optimize the outfit style, a constraint is added based on Equation (7), so the outfits with different styles produce a variety of combinations. The optimization objective function is applied to the adaptive learning rate for each parameter as shown in Equation (8),

$$\operatorname{argmax} (\theta_i) \leftarrow \varepsilon \cdot \left(\sigma (p_{u,i} - \hat{p}_{u,i}) \frac{\theta (p_{u,i} - \hat{p}_{u,i})}{\partial \theta} + \lambda_\theta \theta_i \right). \tag{8}$$

where ε is the learning rate, ε in R, and $\lambda_\theta \theta_i$ is the regularized hyper parameter. $p_{u,i}$ represents the user interacting with the outfits. In the method, the visual style of the item i and j are projected to the style space, if the two items have good compatibility, the distance between f_i and f_j is as far as possible is closer. They belong to the same style and match each other, they have higher compatibility. The algorithm is presented as Algorithm 1.

Algorithm 1 The algorithm for our method

Input: i: item that not rated by the user; j: item that has rated by the user; f_i: features of the product i;
$\quad$ f_j: features of the product i; L_2 regularizer λ_θ, learning rate ε.
Output: Model parameters θ;
$\quad$ 1: **for** every unrated item of current user i **do**
$\quad$ 2: $\quad$ **for** every item that the current user has rated j **do**
$\quad$ 3: $\quad\quad$ $\hat{p}_{u,i}$ = Current user 's rating on item j
$\quad$ 4: $\quad\quad$ Obtain user records that have evaluated both i and j, construct dynamic preference, and simultaneously compute the compatibility of between f_i and f_i according Equation (7).
$\quad$ 5: $\quad\quad$ Compute the sampling probability $\hat{p}_{u,i}$.
$\quad$ 6: $\quad$ **end for**
$\quad$ 7: $\quad$ Update $\hat{p}_{u,i}$ with functional gradient descent according Equation (8);
$\quad$ 8: **end for**
$\quad$ 9: Return the updated model parameters θ

---

4. Experiment

4.1. Datasets and Experiment Data

To verify the performance of the model, and evaluate the model capability and applicability in a variety of real-world scenarios, we collect a large number of data sets from different areas, including user shopping records, and the characteristics of the products. These data sets record various attributes. Among these attributes, visual information is an important factor influencing consumer decisions, as shown in Table 1.

Table 1. Dataset statistics (after preprocessing).

Dataset	Users	Items	Feedback
Amazon Women	99,748	331,173	854,211
Amazon men	34,212	100,654	260,352
Amazon Phones	113,900	192,085	964,477
Epinions.com	19,823	166,526	410,186
Total	267,683	790,438	2,489,226

Amazon. The dataset [18] is widely used in performance test experiments of recommender systems and contain a large number of consumptions of users in different categories. There are 80,000 items and 210,000 transactions in the data set. We take 'Cell Phones and Accessories', and 'Clothing, Shoes and Jewelry' as Phones data sets and Clothing data sets for short. More than 80,000 stylish images were selected from the datasets, each of which was labeled into approximately 50 categories with 1000 descriptive attributes. To get a better training model, we split up the dataset and selected 20 attributes with more balanced data distribution. During the experiment, four attributes are recorded in the dataset, namely user number, item number, rating and scoring time.

Epinions. From a popular consumer review website, Epinions.com [34]. Like the Amazon dataset, it contains the user's purchase history. The purchase record maintains the order of the operations and the praised behavior is used as positive feedback. Each data set is processed by extracting the implicit feedback and attribute characteristics already described. If the user's evaluation record for the item is less than 3, our evaluation of the item is ignored.

4.2. Evaluation Method and Discussion

We split the dataset into training/verifying/testing sets, selecting a random item for each user u to verify $v(u)$, and the remainder of the data set as test T_u. Since all methods use the comparison to optimize the index directly on four training sets, evaluating the forecasted ranking by the metric area used under the curve.

$$AUC = \frac{1}{|\mu|} \sum_{|\mu|} \frac{1}{T_u} \sum_{(i,j) \in T_u} \delta(\widehat{x}_{u,i} > \widehat{x}_{u,j}). \tag{9}$$

To prove the validity of the model and the better performance on the test set, the dimensions of the user and item representations in all methods are set to the same factors and all parameters are tuned. On Amazon datasets, the regularized tuning parameter is 20 in most cases, and BPR-MF [35], FPMC [36], VBPR [18], and CIGAR [37] perform best on area under curve (AUC). For each data set, we report the average AUC (representing all commodities) on the complete test set T and the subset of T, which contains only items with less than 5 positive feedback instances. The results of AUC on different data sets and different algorithms are shown in Table 2, the best performance method on every dataset is boldfaced. On the right we demonstrate the improvement of ours vs the best baselines.

Table 2. *AUC* on the test set. The best performing method in each row is red, and the second-best method in each row is boldfaced.

Algorithm	Fossil		FPMC		BPR-MF		VBPR		CIGAR		Ours	
	Auc@10	Auc@20	Auc@10	Auc@20	Auc@10	Auc@20	Auc@10	Auc@20	Auc@10	Auc@20	Auc@10	Auc@20
Amazon Women	0.6585	0.7255	0.6240	0.7020	0.5992	0.6672	0.6353	0.7163	**0.7284**	**0.7334**	0.7516	0.7562
Amazon Men	**0.7194**	**0.7194**	0.6100	0.7100	0.6766	0.7156	0.6185	0.7185	0.6883	0.7198	0.7563	0.7524
Amazon Phones	0.6484	0.6885	**0.7286**	**0.7418**	0.6259	0.7259	0.5797	0.6797	0.6678	0.6852	0.7584	0.7645
Epinions.com	0.6048	0.7048	-	-	0.6577	0.6774	-	-	**0.7184**	**0.7229**	0.7377	0.7456

4.3. Performance and Quantitative Analysis

It is worth noting that all methods optimize the AUC metrics on the test set, but our method performs better than other methods.

(1) Compared with the VBPR, our method has a greater improvement on the Amazon Women data set. As the sparsity rate increases, the value of AUC increases, which indicates that customers are more likely to buy fashion items on Amazon. However, on the Amazon mobile phone data set, customers have a greater tendency to choose practical items. (2) Compared with methods without visual information, the method with visual information significantly improves the diversity and accuracy of items recommendations. (3) Compared with VBPR and CIGAR, our method can perform better on all datasets, and demonstrates the effectiveness and reliability of learning visual compatibility relationships from external datasets.

In Figure 2, we can observe that AUC increases with increasing epoch and tends to stabilize in the end. Using AUC@10 and AUC@20 respectively, we can observe that the performance of the top 20 commodities recommended is better, and we think that the more items displayed, the more stable the model performance.
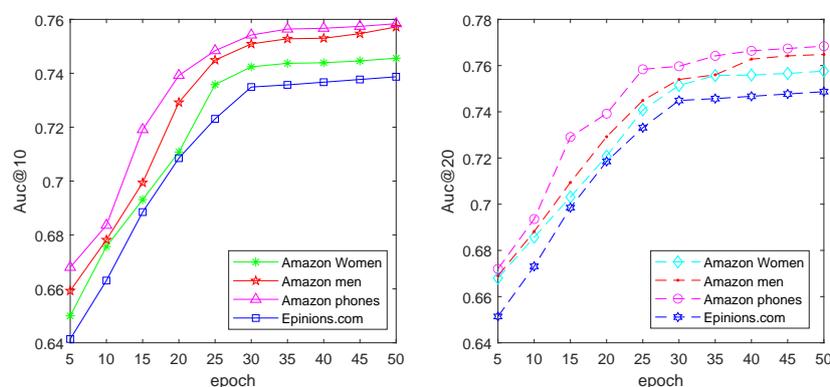

Figure 2. The performance of AUC@k on different epochs. In the experiment, the values of *k* are 10 and 20, respectively.

Training efficiency. AUC is demonstrated by increasing the number of training iterations in the test set. Generally, it takes about 5 hours for our model to converge on the women's dress data set, and the convergence time is longer than BPR-MF and CIGAR.

4.4. Case Study

We select products purchased by Amazon users from 2008 to 2013 for analysis. According to user rating and purchase record, user preferences are found, and under user personalized differences, a variety of clothing and accessories are recommended for matching, as shown in Figure 3. Based on the user's existing product collection, the proposed method recommends to the user articles of clothing that match one another. The collection of images in the four small boxes in Figure 3, include both the user's existing products and the recommended products. A total of eight different styles were selected, each of which was in a different category. The top four for womenswear and the bottom for menswear.

Each row contains a recommended outfit where query images are indicated, and results to some matching items are recommended, the results show that the items in red boxes are more compatible.
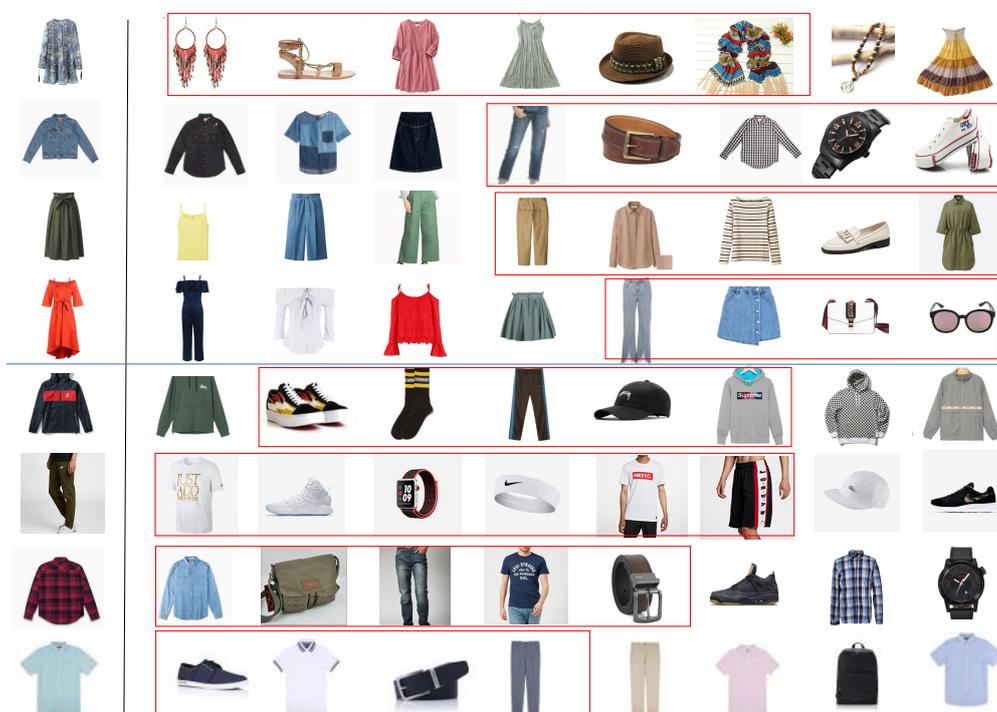

Figure 3. Fashion outfit recommendation is based on query items.

Figure 4 shows that four related apparel items are listed for users to choose and match according to their previous purchase records. The generated clothes can be matched according to gender and style. Above is the category of men, and below is the category of women's clothes that can be matched with the style of men. The same style of clothing and merchandise represented in each of the small squares. We can see that no matter from the casual style, or the style of simplicity, as well as the girl's unique Bohemian style, these styles of internal goods have a high compatibility. The eight outfits are used to represent men's and women's clothes, respectively. Clothing products are matched within the area. Take leisure style as example, the interior of the sport style clothing can be matched with each other. Sports style and leisure style clothing can also be matched with each other.

We select four style outfits for display. In every outfit, the items are compatible with each other and each set contains only one type of item. Because of the limited set of images that are shown in Figure 5, a more comprehensive picture has been used. There are different styles of womenswear in different areas. As shown in Figure 5, there are four style outfits, and the items of the outfits are compatible. Every outfits contain five different items. Based on the user preference, the proposed method generates different style outfits for different scenarios. From which we can draw two conclusions, First, items in different subcategories can produce compatible combinations. Second, the products in each type of matching product collection will be arranged in order according to the comprehensive conditions, and the latent consistency of the products can be learned.
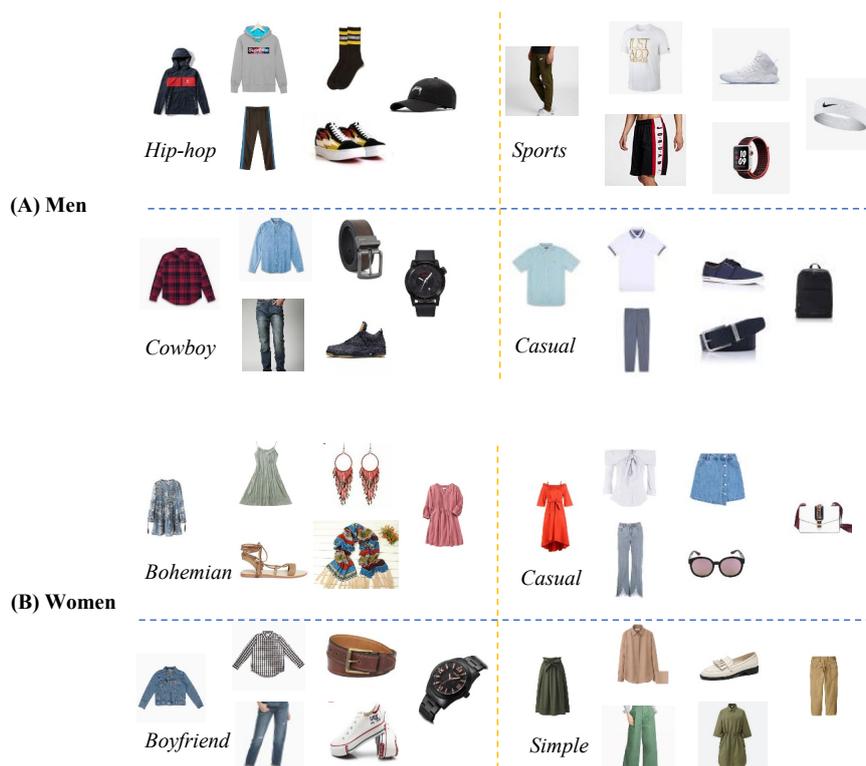
Figure 4. According to our algorithm, we have selected several types of matching clothing, which are divided into men's and girls' clothing, and we have selected the matching clothing products. The above is the men's clothing, and the following is the women's clothing category.

Figure 5. The outfits generated by our algorithm.

We conduct experiments based on Amazon datasets, and propose a clothing matching method based on the user dynamic preference, solving the problems of preferences modeling, network structure, style matching and so on. In addition, we analyze the visual dimension of clothing, personalize the visual style, and use the fewest clothes to match the outfits, to achieve the item optimal combination. The experimental results are consistent with the compatibility and diversity of visual styles. The experiment shows that the visual features of the item integrated into the collaborative filtering algorithm based on user–item interaction have a significant impact on user preferences.

During the experiment, through the classification of clothing products, different kinds of attribute features are extracted. Considering the dependence of hierarchical organizational features on user commodity interaction, we model personalize outfits based on the dynamic changes in user preferences, and to achieve the purpose of recommending more complementary items. To improve the performance

of search results, multiple features must be considered, such as the number of words in the query text, the number of images. The traditional context-aware recommendation system does not take into account the time effect of item recommendations. The proposed method can more accurately understand the user's preferences as well as accurately predict the common preferences of the group in which the user is located.

5. Conclusions

In this paper, we propose a new method for personal style to address personalized behavioral prediction in fashion recommendation. We explicitly model scene-based user preference to infer the item–item relationship from historical user interactions. We have introduced a novel method for the fashion recommendation task, which can estimate the relative weights of each item in user interaction, so as to learn better representations for short-term user interests. The model considers the interaction between users and commodities, and makes full use of the visual relationship of commodities to model the relationship between commodities. The experimental results show that our method has very good performance compared to other methods in the AUC and diversity indicators. Experimental results on the publicly available Amazon data set verify the validity of the model. In addition, we visualized some samples to illustrate the effectiveness of the method. Despite the progress of our work, due to the complex and highly dynamic content of network resources, the user-provided query terms are often short and will lead to ambiguous ambiguity. Relying only on literal matching, it is difficult to understand the search intent contained in user queries. One future direction is to combine structured knowledge and deep natural language understanding to understand user intentions, and adopt a heterogeneous graph approach to solve complex diverse networks.

Author Contributions: Supervision: X.C. (Xiaochun Cheng); Conceptualization, X.C. and A.A.; methodology, J.W.; software, A.A. (Ahed Abugabah); validation: A.A. (Ahed Abugabah); investigation, J.W.; resources, X.C.; Writing—original draft, A.A. (Ahed Abugabah); writing—review and editing: X.C.(Xiaochun Cheng). All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hu, Y.; Yi, X.; Davis, L.S. Collaborative Fashion Recommendation: A Functional Tensor Factorization Approach. In Proceedings of the 23rd ACM international conference on Multimedia, Brisbane, Australia, 26–30 October 2015; pp. 129–138.
2. Veit, A.; Kovacs, B.; Bell, S.; McAuley, J.J.; Bala, K.; Belongie, S.J. Learning Visual Clothing Style with Heterogeneous Dyadic Co-Occurrences. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 4642–4650.
3. Geng, X.; Zhang, H.; Bian, J.; Chua, T. Learning Image and User Features for Recommendation in Social Networks. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 4274–4282.
4. Han, X.; Wu, Z.; Jiang, Y.; Davis, L.S. Learning Fashion Compatibility with Bidirectional LSTMs. In Proceedings of the 25th ACM international conference on Multimedia, Mountain View, CA USA, 23–27 October 2017; pp. 1078–1086.
5. Wang, H.; Li, Z.; Li, Y.; Gupta, B.B.; Choi, C. Visual saliency guided complex image retrieval. *Pattern Recognit. Lett.* **2020**, *130*, 64–72.
6. McAuley, J.J.; Targett, C.; Shi, Q.; van den Hengel, A. Image-Based Recommendations on Styles and Substitutes. In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, Santiago, Chile, 9–13 August 2015; pp. 43–52.
7. Al-Ayyoub, M.; Alandoli, M.N.; Jararweh, Y.; Al-Smadi, M.; Gupta, B. Improving fuzzy C-mean-based community detection in social networks using dynamic parallelism. *Comput. Electr. Eng.* **2019**, *74*, 533–546.

8. Tkalcic, M. Emotions and personality in recommender systems: Tutorial. In Proceedings of the 12th ACM Conference on Recommender Systems, Vancouver, BC, Canada, 2–7 October 2018; pp. 535–536.

9. Verma, S.; Anand, S.; Arora, C.; Rai, A. Diversity in Fashion Recommendation using Semantic Parsing. *arXiv* **2019**, arXiv:1910.08292.

10. Lin, Y.; Ren, P.; Chen, Z.; Ren, Z.; Ma, J.; de Rijke, M. Improving Outfit Recommendation with Co-supervision of Fashion Generation. In Proceedings of the World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; pp. 1095–1105.

11. Wang, X.; He, X.; Feng, F.; Nie, L.; Chua, T. TEM: Tree-enhanced Embedding Model for Explainable Recommendation. In Proceedings of the 2018 World Wide Web Conference, Lyon, France, 23–27 April 2018; pp. 1543–1552.

12. Zhao, Q.; Chen, J.; Chen, M.; Jain, S.; Beutel, A.; Belletti, F.; Chi, E.H. Categorical-attributes-based item classification for recommender systems. In Proceedings of the 12th ACM Conference on Recommender Systems, Vancouver, BC, Canada, 2–7 October 2018; pp. 320–328.

13. Ak, K.E.; Kassim, A.A.; Lim, J.; Tham, J.Y. Learning Attribute Representations With Localization for Flexible Fashion Search. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7708–7717.

14. Simo-Serra, E.; Ishikawa, H. Fashion Style in 128 Floats: Joint Ranking and Classification Using Weak Data for Feature Extraction. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 298–307.

15. Zou, L.; Sun, J.; Gao, M.; Wan, W.; Gupta, B.B. A novel coverless information hiding method based on the average pixel value of the sub-images. *Multimed. Tools Appl.* **2019**, *78*, 7965–7980.

16. Al-Halah, Z.; Stiefelhagen, R.; Grauman, K. Fashion Forward: Forecasting Visual Style in Fashion. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 388–397.

17. Tangseng, P.; Yamaguchi, K.; Okatani, T. Recommending Outfits from Personal Closet. In Proceedings of the WACV 2018: IEEE Winter Conference on Applications of Computer Vision, Lake Tahoe, NV, USA, 12–15 March 2018; pp. 269–277.

18. He, R.; McAuley, J. VBPR: Visual Bayesian Personalized Ranking from Implicit Feedback. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; pp. 144–150.

19. Xia, M.; Li, T.; Liu, L.; Xu, L.; Gao, S.; de Silva, C.W. Remaining useful life prediction of rotating machinery using hierarchical deep neural network. In Proceedings of the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada, 5–8 October 2017; pp. 2778–2783.

20. Chen, J.; Zhang, H.; He, X.; Nie, L.; Liu, W.; Chua, T. Attentive Collaborative Filtering: Multimedia Recommendation with Item- and Component-Level Attention. In Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval, Tokyo, Japan, 7–11 August 2017; pp. 335–344.

21. Rendle, S.; Freudenthaler, C.; Gantner, Z.; Schmidt-Thieme, L. BPR: Bayesian Personalized Ranking from Implicit Feedback. *arXiv* **2012**, arXiv:1205.2618.

22. He, R.; McAuley, J. Fusing Similarity Models with Markov Chains for Sparse Sequential Recommendation. In Proceedings of the 2016 IEEE 16th International Conference on Data Mining (ICDM), Barcelona, Spain, 12–15 December 2016; pp. 191–200.

23. Hidasi, B.; Tikk, D. General factorization framework for context-aware recommendations. *Data Min. Knowl. Discov.* **2016**, *30*, 342–371.

24. LeCun, Y.; Bengio, Y.; Hinton, G.E. Deep learning. *Nature* **2015**, *521*, 436–444.

25. Tang, J.; Wang, K. Personalized Top-N Sequential Recommendation via Convolutional Sequence Embedding. In Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, Marina Del Rey, CA, USA, 5–9 February 2018; pp. 565–573.

26. You, J.; Wang, Y.; Pal, A.; Eksombatchai, P.; Rosenberg, C.; Leskovec, J. Hierarchical Temporal Convolutional Networks for Dynamic Recommender Systems. In Proceedings of the World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; pp. 2236–2246.

27. Kang, W.; McAuley, J. Self-Attentive Sequential Recommendation. In Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM), Singapore, 17–20 November 2018; pp. 197–206.

28. Hidasi, B.; Karatzoglou, A.; Baltrunas, L.; Tikk, D. Session-based Recommendations with Recurrent Neural Networks. In Proceedings of the 4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, 2–4 May 2016.

29. He, Y.; Zhang, Y.; Liu, W.; Caverlee, J. Consistency-Aware Recommendation for User-Generated Item List Continuation. In Proceedings of the 13th International Conference on Web Search and Data Mining, Houston, TX, USA, 3–7 February 2020; pp. 250–258.

30. Jing, Y.; Zhang, X.; Wu, L.; Wang, J.; Feng, Z.; Wang, D. Recommendation on Flickr by combining community user ratings and item importance. In Proceedings of the 2014 IEEE International Conference on Multimedia and Expo (ICME), Chengdu, China, 14–18 July 2014; pp. 1–6.

31. Iftikhar, Z.; Jangsher, S.; Qureshi, H.K.; Aloqaily, M. Resource Efficient Allocation and RRH Placement for Backhaul of Moving Small Cells. *IEEE Access* **2019**, *7*, 47379–47389.

32. Farouk, A.; Zhen, D. Big data analysis techniques for intelligent systems. *J. Intell. Fuzzy Syst.* **2019**, *37*, 3067–3071.

33. Liu, W.; Wang, J.; Sangaiah, A.K.; Yin, J. Dynamic metric embedding model for point-of-interest prediction. *Future Gener. Comput. Syst.* **2018**, *83*, 183–192.

34. Zhao, T.; McAuley, J.; King, I. Leveraging Social Connections to Improve Personalized Ranking for Collaborative Filtering. In Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, Shanghai, China, 3–7 November 2014; pp. 261–270.

35. Rendle, S.; Freudenthaler, C.; Gantner, Z.; Schmidt-Thieme, L. Bayesian Personalized Ranking from Implicit Feedback. In Proceedings of the UAI, Montreal, QC, Canada, 18–21 June 2009; pp. 452–461.

36. Rendle, S.; Freudenthaler, C.; Schmidt-Thieme, L. Factorizing personalized Markov chains for next-basket recommendation. In Proceedings of the 19th International Conference on World Wide Web, Raleigh, NC, USA, 26–30 April 2010; pp. 811–820.

37. Kang, W.; McAuley, J.J. Candidate Generation with Binary Codes for Large-Scale Top-N Recommendation.

In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing, China, 3–7 November 2019; Zhu, W.; Tao, D.; Cheng, X.; Cui, P.; Rundensteiner, E.A.; Carmel, D.; He, Q.; Yu, J.X., Eds.; ACM: New York, NY, USA, 2019; pp. 1523–1532.