ARTICLE

CNN Recommender Systeme For Fashion

Chaimaa Bouabd* and Hicham MACHAY

Master IAII, FST, Cadi Ayyad University, Marrakech, Morocco

*Corresponding author: chaimabouabd2000@gmail.com

Abstract

Recommender systems (RS) are designed to tackle the contemporary challenge of information overload in e-commerce platforms and other data centric online services. These systems play a pivotal role in aiding users to navigate and make the most of the vast information available within the system. They achieve this by integrating both implicit and explicit data, sourced from internal e-commerce systems and user interactions. In the realm of modern product catalogs, the incorporation of images serves to provide a rapid visual overview, effectively transforming potential buyers into satisfied customers.

Recognizing that the majority of e-commerce establishments utilize product images to engage users' visual preferences and drive purchasing decisions, this paper introduces an image-based Content-Based Filtering approach for RS. This innovative approach leverages deep learning techniques and is rooted in both the item's description and the user's preferred choices profile.

In a content-based recommendation system, item features are utilized to characterize products, and a user profile is established to articulate the user's preferences. Essentially, the algorithms aim to recommend products that align with the user's demonstrated interests. For this study, we employ the VGG16 Pretrained model, a convolutional neural network (CNN), to extract features from product images.

Despite the absence of explicit user profile data, we utilize K-Nearest Neighbors algorithms to recommend products with visually similar features. This mirrors the familiar concept found on shopping websites: "Products that are similar to this." This methodology demonstrates a heightened level of accuracy in product recommendations on e-commerce platforms compared to traditional approaches.

Keywords: Recommendation Systems, Fashion Outfit Recommendation, Deep Learning, Image-based recommender systems, E-commerce, Convolutional neural network, Fashion recommendations, Personalized Fashion Recommendation

1. Introduction

In the ever-evolving landscape of fashion, our paper addresses the demand for personalized recommendations driven by cutting-edge technologies. Exploring the intersection of fashion and artificial intelligence, we recognize the limitations of conventional methods in capturing individual taste nuances. This prompts us to grapple with the pivotal challenge of providing recommendations that genuinely resonate with individual sensibilities.

Navigating the diverse fashion landscape, our research acknowledges the imperative to bridge the gap using advanced technologies. Existing methods, particularly collaborative filtering, fall short in capturing the richness of visual elements crucial to fashion preferences. This recognition forms the foundation for our innovative approach: "Product Recommendations: Visually Similar Content Filtering using KNNs and VGG-16 (CNN)." This model leverages Content-Based Filtering, employing K-Nearest Neighbours (KNN) and integrating VGG-16, a Convolutional Neural Network. Our proposed approach excels in suggesting items beyond conventional limitations, aligning profoundly with unique fashion sensibilities.

As we embark on this exploration, our aim is to illuminate the inner workings of our Content-Based

Filtering model. Delving into the intricacies of integrating KNNs and VGG-16, we highlight how this approach not only addresses conventional challenges but propels fashion recommendation systems toward enhanced personalization and accuracy. Join us on this comprehensive journey through the fusion of fashion and artificial intelligence, where we unravel the threads of innovation to weave a tapestry of style recommendations as unique as the individuals who wear them.

In the forthcoming sections of our paper, we will delve into related works, exploring existing methodologies in the field. We will then outline our proposed approach, "Product Recommendations: Visually Similar Content Filtering using KNNs and VGG-16 (CNN)," emphasizing the integration of K-Nearest Neighbours and Convolutional Neural Networks. Subsequently, we will present and discuss the results, offering insights into how our model enhances fashion recommendation systems in terms of personalization and accuracy.

2. RELATED WORK

2.1 Collaborative Filtering in Fashion Recommendation:

Early approaches to fashion recommendation often relied on collaborative filtering methods that analyze user-item interactions. Explore seminal works in collaborative filtering for fashion, emphasizing the use of user behavior and preferences without relying on image data.

2.2 Content-Based Fashion Recommendation:

Content-based recommendation systems leverage item features to make suggestions. Investigate research in the fashion domain that utilizes content-based methods, emphasizing textual information, metadata, and non-image features extracted from clothing items.

2.3 Hybrid Recommendation Systems:

Hybrid recommendation systems integrate collaborative filtering and content-based methods to leverage both user behavior and item features. Examine studies that have successfully combined these techniques in the context of fashion, focusing on non-image data.

2.4 Rule-Based Systems for Fashion:

Traditional rule-based recommendation systems employ predefined rules to suggest items based on non-image attributes. Survey literature on rule-based approaches for fashion recommendation, discussing how rules are derived and applied without relying on visual information.

2.5 Matrix Factorization Techniques:

Matrix factorization is a popular collaborative filtering technique that decomposes the user-item interaction matrix. Review studies applying matrix factorization in fashion recommendation and highlight variations such as Singular Value Decomposition (SVD) and Alternating Least Squares (ALS) without using image data.

2.6 Graph-Based Recommendation Models:

Investigate the use of graph-based models for fashion recommendation without incorporating image information. Discuss how graph-based methods represent item relationships based on non-image attributes, providing insights into item connections.

2.7 Ensemble Methods in Fashion Recommendation:

Ensemble methods combine predictions from multiple models to improve recommendation accuracy, even without considering image features. Examine how ensemble techniques have been applied to fashion recommendation, emphasizing the fusion of different non-image-based models.

2.8 Clustering Approaches for Fashion:

Explore clustering algorithms applied to group similar fashion items or users based on non-image features. Discuss the suitability of clustering techniques in capturing patterns in fashion datasets without relying on visual information.

2.9 Demographic-Based Recommendation:

Investigate research that incorporates demographic information, such as age, gender, and location, into fashion recommendation systems without using image data. Analyze how demographic factors contribute to personalized and context-aware recommendations based on non-visual features.

3. Methodology

3.1 Method overview

- A recommender system, or a recommendation system, is a subclass of information filtering system that provide suggestions for items that are most pertinent to a particular user.Typically, the suggestions refer to various decision-making processes, such as what product to purchase, what music to listen to, or what online news to read. Recommender systems are particularly useful when an individual needs to choose an item from a potentially overwhelming number of items that a service may offer.

- Recommender Systems can be broadly classified into 3 types
  - Collaborative Filtering
  - Content-Based Filtering
  - Hybrid

- This notebook will demonstrate the Content-Based filtering, which are based on the description of an item and a profile of the user's preferred choices. In a content-based recommendation system, features are used to describe the items, besides, a user profile is built to state the type of item this user likes. In other words, the algorithms try to recommend products that are similar to the ones that a user has liked in the past.

- Although we do not have any kind of user profile data, we will use K-Nearest Neighbours algorithms to recommend products which have visually similar features, such as the ones you see in shopping websites e.g. "Products that are similar to this"

3.2 Dataset

Data used in this work is **Fashion Product Images Dataset** from kaggle The growing e-commerce industry presents us with a large dataset waiting to be scraped and researched upon. In addition to professionally shot high resolution product images, we also have multiple label attributes describing the product which was manually entered while cataloging. To add to this, we also have descriptive text that comments on the product characteristics. Each product is identified by an ID like 42431. You will find a map to all the products in **styles.csv**. From here, you can fetch the image for this product from images/42431.jpg and the complete metadata from **styles/42431.json**. To get started easily, we also have exposed some of the key product categories and it's display name in styles.csv. If this dataset is too large, you can start with a smaller (280MB) version https://www.kaggle.com/paramaggarwal/fashion-product-images-small

3.3 EDA and Visualization

In our dataset, we have two CSV files: "images.csv," which contains the ID and link of fashion products, and "styles.csv," which contains metadata for each product. We initiate the data preprocessing by merging these two datasets using a unique ID. Subsequently, we filter out products for which images are not present. After this step, we check for the existence of any null values. Following the data cleaning process, we visualize the dataset and remove unnecessary columns such as "productDisplayName," "link," and "file found."

While we won't follow a traditional training-validation approach, we still perform some steps on the training set and apply the learned weights to the validation set. The data is shuffled randomly, and an 80-20 split is applied, allocating 80 % of the data for training and 20 % for validation.

Finally, we incorporate a Data Generator for various reasons, including Memory Efficiency, Parallel Processing, and On-the-Fly Augmentation, among others. This approach enhances the training process by efficiently managing large datasets, processing data in parallel, and dynamically augmenting data during training for improved model generalization.

3.4 Image-based feature extraction

In the process of extracting features from images, leveraging the capabilities of pre-trained Convolutional Neural Networks (CNNs) has become a common practice. One popular architecture for feature extraction is VGG16. By utilizing a pre-trained VGG16 model, we tap into the hierarchical representation of features learned from vast datasets during its training phase. To extract image features, we focus on the final convolutional block of the VGG16 network. Employing global average pooling on this block allows us to aggregate and condense the spatial information across each feature map, producing a compact representation of the image's distinctive features. This approach not only captures meaningful patterns but also facilitates a reduction in the dimensionality of the feature space. Leveraging pre-trained models like VGG16 for feature extraction proves valuable in various applications, from image recognition to transfer learning tasks.

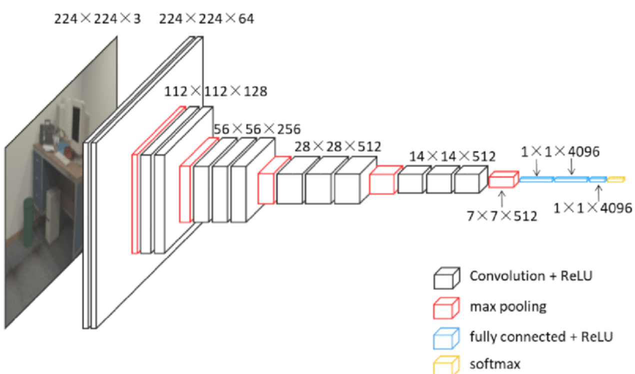

Figure 1. VGG16 Architecture.

3.5 Dimensionality Reduction

this process involves transforming high-dimensional data into a lower-dimensional representation while retaining essential information. In this specific case, the original feature space with 512 dimensions is reduced to a more concise form with only 313 dimensions. The significance of this reduction lies in the fact that the first 313 principal components encapsulate approximately 99% of the variance present in the original data. By retaining such a substantial percentage of variance, we achieve a substantial reduction in dimensionality without sacrificing critical information. This streamlined representation, composed of the most informative components, not only facilitates a more efficient computational process but also enhances our ability to discern and interpret essential patterns within the data.

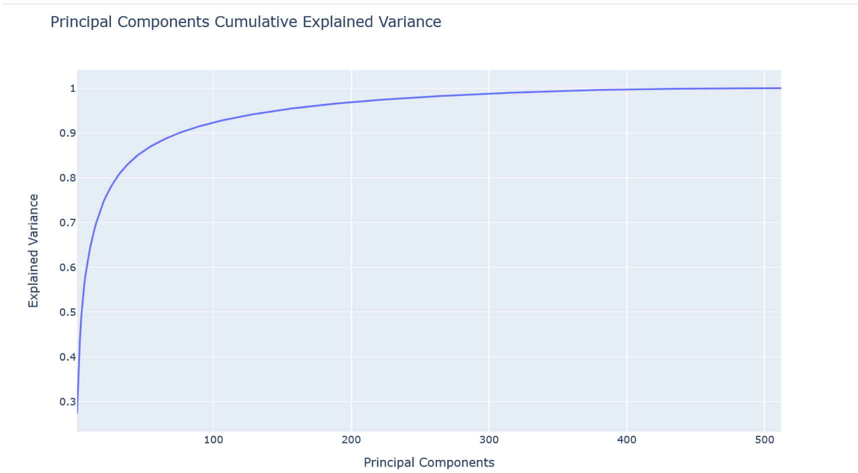

Figure 2. PCA .

3.6 K-Nearest Neighbours

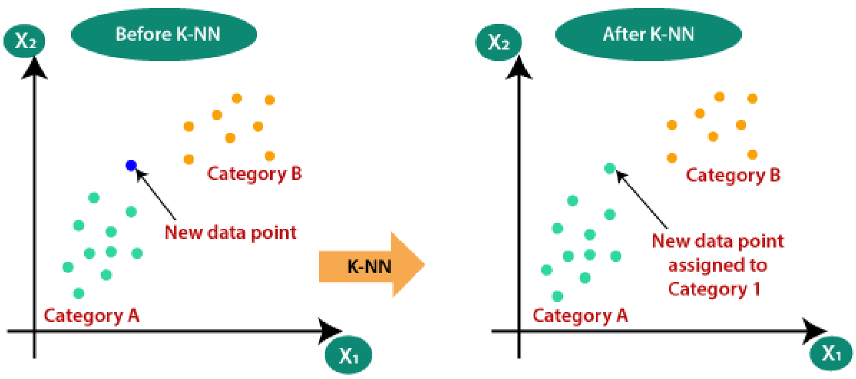

Figure 3. knn schema.

In the context of the K-Nearest Neighbours (KNN) algorithm, a thoughtful approach to product recommendation involves setting the parameter k to 6. This choice allows us to recommend the 6 most visually similar products to a given query product. The rationale behind selecting k=6 is rooted in the understanding that the first product among the recommended 6 will be the query product itself. By examining the subsequent 5 products in the list, we can identify those that exhibit the greatest visual resemblance based on the raw extracted features from a pre-trained neural network.

It's important to note that, unlike traditional classification prediction where the predict method is often employed, the KNN method is used to find the k (in this case, 6) nearest neighbors. These neighbors represent products with visually similar content, making it a valuable technique for content-based recommendation systems.

4. Discussions

The evaluation of such recommendation techniques typically involves metrics such as the hit-rate. The hit-rate serves as an indicator of the system's effectiveness by measuring whether the user ultimately purchased the product that was recommended to them. This evaluation metric provides valuable insights into the practical success of the recommendation system in delivering products that align with the user's preferences and needs.

Figure 4. Results.

Figure 5. Results.

- The results can definitely improved by creating data of similar styled or branded products
- A siamese network can then be trained to create image features where similarly styled or branded products are in closer proximity

Table 1. Results of the comparison between VGG16 and ResNet50.

bla bla	VGG16	ResNet50
Mean Squared Error (Train)	143.8545940621922	248.40970873786407
Root Mean Squared Error (Train)	11.99393988905198	15.761018645311733
Mean Squared Error (Test)	589.4747861323729	778.7179198559207
Root Mean Squared Error (Test)	21.8731170798296	27.905517731372065

Due to the absence of user ratings for evaluating the performance of our system, we adopted an alternative approach. Initially, we attempted to label our data by utilizing the 'articleType' as our designated label. This labeled data was then transformed into categorical format, and the extracted features served as input for our evaluation. In this comparative analysis, we employed the k-Nearest Neighbors (KNN) algorithm using features extracted by both the VGG16 and ResNet50 architectures. The implementation of this approach revealed that VGG16 successfully captured more meaningful patterns than the ResNet50 architecture. This outcome suggests that, in this context, VGG16 excelled in extracting distinctive features that contribute to the overall effectiveness of the system.

5. Conclusion

Acknowledgement

Include your acknowledgement in this section.

Author contributions

If the paper has more than one author, the CRediT section must be included. See example usage on https://casrai.org/credit/

- First Author: Conceptualization, Methodology, Software, Writing- Original draft
- Second Author: Data curation, Writing- Original draft
- Third Author: Visualization, Investigation

References

[1] Davidson, J et al. (2010). 'The youtube video recommendation system. In Proceeding of the 4th ACM Conference Recommender Systems, pp. 293-296. IEEE

[2] Shan, ZP, Y Q Lei, NASM: Nonlinearly attentive similarity model for recommendation system via locally attentive embedding. IEEE Access, 7, 70689-70700.

[3] Li, J-B, SY Lin, Y-H Hsu and Y-C Huang (2019). Implementation of an alternating least model based collaborative fillering movie recommendation system on hadoop and spark platforms. BWCCA, 25, 237-249.

[4] Pan, R, Y Zhou, B Cao, NN Liu, R Lukose, M Scholz and Q Yang (2008). One-class collaborative filtering. In Proceedings of the IEEE International Conference on Data Mining (ICDM), pp. 1-10.

[5] Zhang, J, Y Wang, 2 Yuan and Q Jin (2018), Personalized real-time movie recommendation system: Practical prototype and evaluation. Tsinghua Science, 25, 180-191.

[6] Liu, J, D Wang and Y Ding (2017). PHD: A probabilistic model of. hybrid deep collaborative filtering for recommender systems. In Proceedings of the 9th Asian Conference on Machine Learning, PMLR, Vol. 77, pp. 224 239.

[7] Maxwell Harper, F and JA Konstan (2015). The MovieLens datasets: History and context. ACM Transactions on Interactive Intelligent Systems (TiS), 5, 1-19.

[8] Yongsheng, L, H Qu, W Chen and SM Hasan Mahmud (2019)., An efticient deep learning model to infer user demographic information from ratings. IBEE Access, 7, 53125-53135.

[9] Meira, D, J Viterbo and F Bernardini (2018). An experimental analysis on scalable implementations of the alternating least squares algorithm, IeEE, 15, 351-359. Mu, R (2018). A survey of recommender systems based on deep learning. IEEE Access, 6, 69009 69022.

[10] Mu, R (2018). A survey of recommender systems based on deep learning. IEEE Access, 6, 69009 69022.

[11] Ning, X and G Karypis (2011). SLIM: Sparse linear methods for top-N recommender systems. ICDM, pp. 1-10. DO1:10.1109/ICDM.2011.134.

[12] H. Zhao, C. Shi, X. Li, P. Zhao, and J. Fang, "Session-Based Collaborative Filtering with Hierarchical Graph Convolutional Networks," in Proceedings of the 14th ACM International Conference on Web Search and Data Mining, 2021, pp. 14-22.

[13] Lops, P., de Gemmis, M. and Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In Recommender systems handbook (pp. 73-105). Springer US.

[14] N. Agarwal, M. Patel, and R. Garg, "An Efficient and Scalable Algorithm for Content-Based Movie Recommendation using Matrix Factorization with Block-Wise Feature Selection," Journal of King Saud University - Computer and Information Sciences, vol. 33, no. 3, pp. 295-307, 2021.

[15] Robin Burke and Kristian Lum (2006).Improving Recommendation Accuracy by Combining Collaborative Filtering and Content-Based Filtering. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

[16] Zhao, X., Zhang, J. and Zhang, H. (2021). A Hybrid Matrix Factorization and Neural Network Model for Personalized Recommendations. IEEE Access, 9, 12835-12844. doi: 10.1109/ACCESS.2021.3058898