Enhancing Interior Design

A Furniture Recommender System Integrating Deep Learning

Laksshay Sehrawat 2021UCA1802
Vanshika Gupta 2021UCA1805
Amrita Soney 2021UCA1831

ABSTRACT

Recommender systems are a subclass of information filtering systems. These systems are specialized software components, which usually make part of a larger software system, but can also be standalone tools. A recommender system's main goal is to provide the user software suggestions for items that can be useful. The suggestions are related to different decision-making mechanisms, different techniques, such as, what product to buy, what movie to watch, or what vacation to reserve. In the context of recommender systems, the general term "item" refers to what the system is actually recommending to its users. The paper presents the development of a recommendation system, capable of making item suggestions, based on image provided by the user and interaction data, using an image- based CNN algorithm.

PROBLEM STATEMENT

Customer Perspective

Overwhelming Variety

The vast furniture selection leaves customers uncertain about pieces that match their design preferences, lacking efficient guidance tools.

Absence of Image - Based Recommendations

Existing recommendations rely on history, not the user's room image, leading to suboptimal matches for the space.

Inadequate Visualization

Current online shopping experiences struggle to accurately show how furniture fits a room's design theme, leading to post-purchase dissatisfaction.

INTRODUCTION

The main objective of our recommender systems in a furniture company is to narrow down a vast catalog to assist customers in finding exactly the products they want. Most recommendation algorithms (Collaborative filtering and Content-based filtering) that are used in e-commerce companies leverage customer's browsing history. However, the uncertainty, diversity, and timelines of each customer's profile, as well as the absence of new customers' history, make it challenging for such algorithms to be robust to all customers. Hence, we have implemented the newest method for aiding customers. This model leverages an image-based model (CNN) to understand compatibility from product imagery, thereby mimicking the way professional interior designers match pieces of furniture together and eliminating the cold start problem in the process.

LITERATURE REVIEW

Name	An Intelligent Recommendation System Model based on Style for Virtual Home Furnishing in Three-dimensional Scene
Authors	Yan Wang and Hengyu Wang, Xirui Li
Summary	The recommendation system operates on the concept of unified style to establish a 3D furniture recommendation system. This involves categorizing furniture into clusters, identifying user preferences through physiological signal collection and user information modeling, and suggesting other furniture items with a similar style. By conducting model retrieval within the same clusters, the overall search time is reduced, leading to more practical and efficient results, thereby enabling dynamic recommendations. Recognizing the enhanced information representation and content conveyance capabilities of 3D models, especially in line with Marr D's theory of three-dimensional visual environment information, furniture products displayed in 3D virtual scenes with user interaction become more user-friendly and memorable. This aligns with user observation habits, offering a tangible, interactive environment for users to experience products.

LITERATURE REVIEW

Name	Personalized fashion recommender system with image-based neural networks
Authors	M Sridevi et al
Summary	In this paper, they proposed a personalized Fashion Recommender system that generates recommendations for the user based on an input given. Unlike the conventional systems that rely on the user's previous purchases and history, this project aims at using an image of a product given as input by the user to generate recommendations that are similar to that. We use neural networks to process the images from the Deep Fashion data set and a nearest neighbor-backed recommender to generate the final recommendations. In this paper, the author has presented a novel framework for fashion recommendation that is driven by data, visually related, and simple effective recommendation systems for generating fashion product images. The proposed approach uses a two -stage phase. Initially, their proposed approach extracts the features of the image using CNN classifier i.e. , for instance allowing the customers to upload any random fashion image from any E-commerce website and later generating similar images to the uploaded image based on the features and texture of the input image.

LITERATURE REVIEW

Name	Furniture style compatibility recommendation with cross-class triplet loss
Authors	Tse-Yu Pan, Yi-Zhu Dai, Min-Chun Hu, Wen-Huang Cheng
Summary	This paper introduces an approach for assessing the compatibility of 3D furniture styles across different classes. The method is built upon the utilization of Triplet CNN to extract more representative features that effectively capture the style compatibility between pairs of 3D furniture models[3]. To validate it, a dataset containing 420 textured 3D furniture models was collected and assessed by a group of raters recruited from Amazon Mechanical Turk (AMT) to evaluate the comparative suitability of paired models within the dataset. The proposed method was trained and evaluated based on three datasets, including the collected textured dataset and two additional non-textured datasets. Experimental results indicate that the method outperforms the state-of-the-art approach, which learned a metric using pre-extracted geometric features. Furthermore, the method is utilized to create a furniture recommendation system that can assist potential users in interior design decisions.

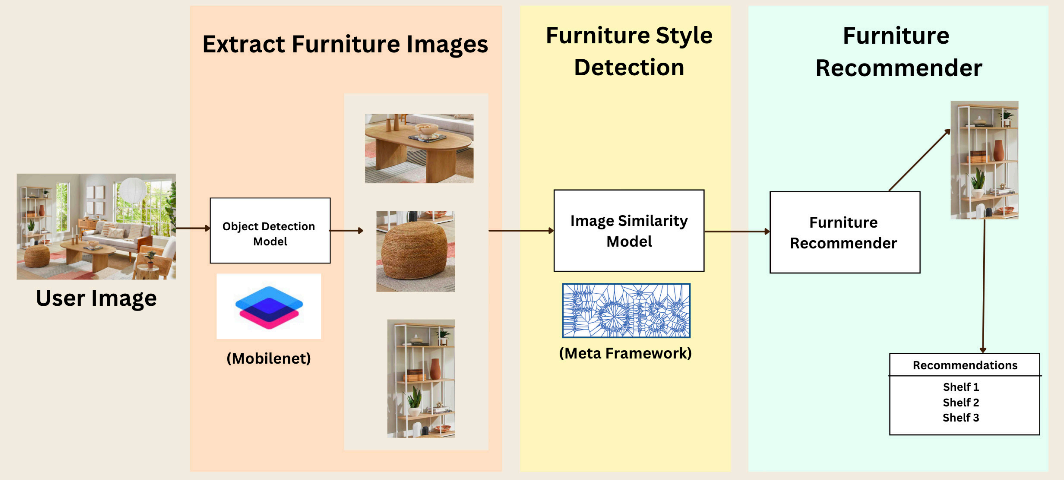
LITERATURE REVIEW

Name	A Multimodal Recommender System for Large-scale Assortment Generation in E-commerce
Authors	Murium Iqbal, Adair Kovac, Kamelia Aryafar
Summary	The assortment recommender system leverages product images to establish visually coherent trends within Overstock's product offerings. Two variants are introduced: a visual-only version and a multimodal version. The visual-only variant generates a representation of product images by applying thresholding to activations from specific layers of a pre-trained deep residual neural network, Resnet-50. It subsequently employs topic modeling (LDA) on these product image representations to create visual trends among Overstock products. A greedy approach is then proposed, both with and without budget constraints, to formulate assortment recommendations based on seed items that maximize the visual compatibility of the set. The multimodal variant utilizes not only image representation but also text-based product attributes. This variant harnesses Polylingual LDA (PolyLDA) to establish trends based on two modalities: images and text. The experimental findings demonstrate that the incorporation of both image and text data results in a more cohesive visual style compared to using only images, thereby enhancing user engagement metrics within the recommendations module.

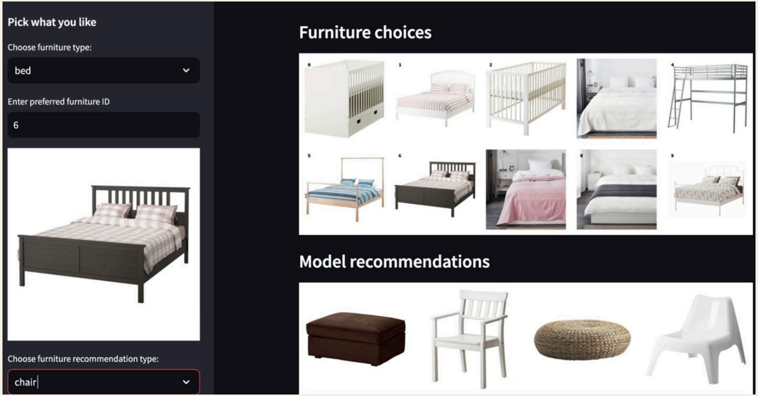
LITERATURE REVIEW

Name	A Machine Learning based Recommendation System for furniture selection
Authors	Mar´ıa Isabel Manresa Rom´an, Javier Ruiz Hidalgo Advisor, Pol Albacar Fern´andez
Summary	The idea was a system that aimed to identify different pieces of furniture in a room and tell the user where to buy or find them. They have built a bed recommender that has been built with its own database as well as a detector, a classifier, and an Image retrieval block. The system retrieves similar and relevant images to the user with all the information needed and also relies on a User interface that makes it easier for a user to interact with it. They have a database with 386 images separated into 3 categories by a classifier that has an accuracy of 95.41\%. A Web Scraping algorithm has been developed for a specific site, in this case, Ikea. A user-friendly interface has an efficient and well-structured code. Their future work is to improve the performance of the whole system. A new classifier with more capacity and ability to choose between master and individual beds needs to be implemented. Right now the recommender only works for beds, but it would be useful to have bedding and to extend the possible recommendations to all bedroom furniture and even for other spaces.

MODELLING OVERVIEW



DEMONSTRATION

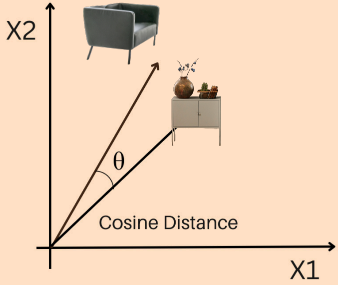


Github link-

https://github.com/Laksshay-Sehrawat/Deep-learning-recommender

MODELLING OVERVIEW

1. Remove the top layer of the trained model for prediction
2. Pre-process images on the Model
3. Collect vector output for all IKEA products and user image
4. Compare Vectors using Cosine Similarity



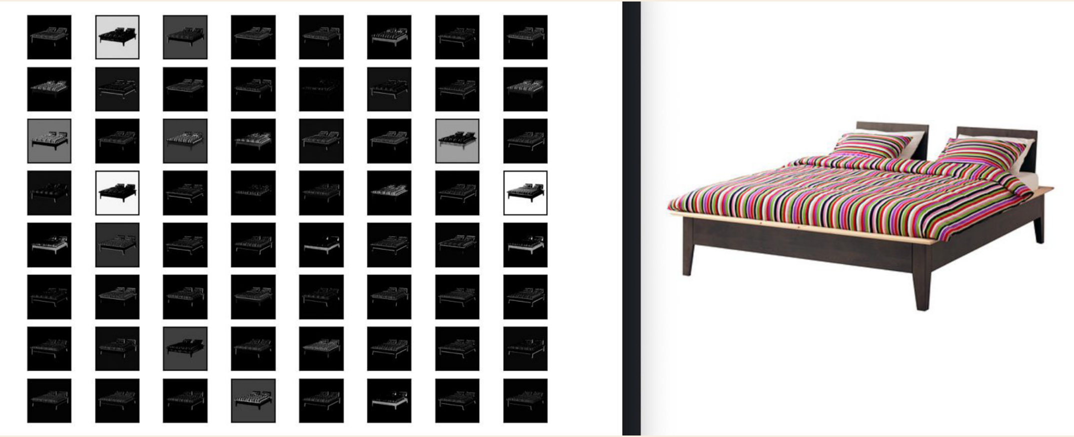
OBJECT DETECTION

Object Detection model used is **MobileNet v2**, is designed for detecting furniture within room images and extracting it.

VGG16 vs VGG19

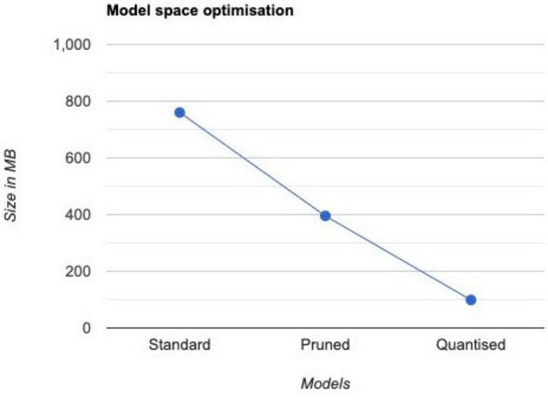
VGG16 and VGG19 are both **convolutional neural network (CNN)** architectures that belong to the VGGNet family. They were introduced by the Visual Graphics Group (VGG) at Oxford University and are known for their simplicity and effectiveness in image classification tasks. The main difference between the two lies in their depth, specifically in the number of convolutional layers. **VGG16 comprises 16 layers (13 convolutional and 3 fully connected layers) and VGG19 extends to 19 layers (16 convolutional and 3 fully connected layers).** VGG19's deeper architecture implies a higher number of parameters, potentially enabling it to capture more intricate features in data, but also demanding greater computational resources compared to VGG16. While VGG16 is computationally less intensive and requires fewer parameters, making it suitable for scenarios with limited computational resources and a smaller dataset, VGG19 may be preferred when tackling tasks that benefit from a more complex model, particularly when working with larger datasets. The choice between VGG16 and VGG19 ultimately hinges on the specific requirements of the task, available resources, and the nature of the dataset being used for training.

MODELING LAYERS



CNN Layers visualization

MODEL SPACE OPTIMIZATION



General:

Kind: Document
Size: 75,99,98,728 bytes (760 MB on disk)

General:

Kind: Document
Size: 39,52,89,880 bytes (395.3 MB on disk)

General:

Kind: Document
Size: 9,89,08,680 bytes (98.9 MB on disk)

We optimized a standard 760 MB model via pruning, reducing it to 395.3 MB, and further applied quantization, resulting in a highly compact 98.9 MB transfer learning model, effectively addressing computational constraints.

EVALUATION

Ideally, rigorous A/B testing was planned to assess the accuracy of its product recommendations. However, due to time and resource constraints, conducting such testing for Pair wasn't feasible. Consequently, an alternative approach was devised. The IKEA image dataset, which included item-to-room image and text information, presented an opportunity to evaluate our model's recommendations differently. The dataset featured curated furniture groupings in various room scenes, assumed to align with Swedish aesthetics.

These furniture arrangements, believed to be carefully selected by IKEA interior designers, served as a benchmark to measure the similarity between Pair's recommendations and IKEA interior designs. Specifically, the metric employed for this evaluation was hit ratio at n (HR@n). A "hit" was registered when one of Pair's top-n recommendations matched the IKEA interior designs. HR@n was calculated as the number of recommendations with a correct top-n match divided by the total number of recommendations issued. A high hit ratio, close to 1, indicated strong agreement between Pair's recommendations and IKEA interior designs, while a ratio of 0 suggested poor agreement. For example, if the HR@5 score was 0.5, it meant that in 50% of all furniture queries, at least one of our top-5 recommendations correctly matched one of the IKEA interior designs.

```
==> Query time: 53.184 ms
102.335.32.jpg  HR@n: None        len(results_files_all): 23
==> Query time: 43.229 ms
500.395.52.jpg  HR@n: 2           len(results_files_all): 32
==> Query time: 46.165 ms
102.051.81.jpg  HR@n: None        len(results_files_all): 20
==> Query time: 38.189 ms
701.032.50.jpg  HR@n: None        len(results_files_all): 14
==> Query time: 69.313 ms
902.396.67.jpg  HR@n: None        len(results_files_all): 31
==> Query time: 33.393 ms
490.904.81.jpg  HR@n: None        len(results_files_all): 23
==> Query time: 59.231 ms
502.954.72.jpg  HR@n: 3           len(results_files_all): 31
HR@5: 1340 (67.36%)
HR@10: 1420 (71.09%)
HR@20: 1440 (72.27%)
```

CONCLUSION

The image-based product recommender is a robust solution leveraging CNN-powered design feature extraction. Deployed as a Dockerized Streamlit container in the cloud, it excels in making cross-category recommendations based on design similarities. The live demo showcases its capability and skill in making accurate and effective recommendations based on design features or similarities. Evaluating it using IKEA room scenes, the hit ratio at 'n' emerged as a compelling metric, affirming its capability to align recommendations closely with IKEA interior designs. A transfer learning pipeline was established to further refine feature maps, enabling more precise detection of furniture pieces that are personalized and style-aligned product recommendations, continually evolving to enhance user experience.

FUTURE GOALS

1. Introducing user feedback mechanisms that will enable personalized recommendations and learning from individual preferences. Exploring advanced deep learning techniques, like attention mechanisms or generative models, could refine its understanding of design elements.

2. Integrating product metadata and contextual information, such as material, size, or room context, will further tailor recommendations.

FUTURE GOALS

3. Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation - We can present a new approach for the "personalization" of text-to-image diffusion models. Given as input just a few images of a subject, we fine-tune a pre-trained text-to-image model such that it learns to bind a unique identifier with that specific subject. Once the subject is embedded in the output domain of the model, the unique identifier can be used to synthesize novel photorealistic images of the subject contextualized in different scenes.

REFERENCES

1. Yan Wang and Hengyu Wang, Xirui Li, An Intelligent Recommendation System Model based on Style for Virtual Home Furnishing in Three-dimensional Scene, International Symposium on Computational and Business Intelligence 2013
2. Mami Mori, Jason Orlosky, Kiyoshi Kiyokawa, Haruo Takemura, IEEE International Symposium on Mixed and Augmented Reality Adjunct Proceedings 2016
3. Tse-Yu Pan, Yi-Zhu Dai, Min-Chun Hu, Wen-Huang Cheng, Furniture style compatibility recommendation with cross-class triplet loss, Multimed Tools Appl (2019) 78:2645–2665
4. Murium Iqbal, Adair Kovac, Kamelia Aryafar, A Multimodal Recommender System for Large-scale Assortment Generation in E-commerce, ACM's Special Interest Group on Information Retrieval (SIGIR) workshops, 2018
5. Mar´ıa Isabel Manresa Rom´an, A Machine Learning based Recommendation System for furniture selection, (UPC): Javier Ruiz Hidalgo Advisor, Pol Albacar Fern´andez Barcelona, 2022

REFERENCES

6. L.H. Li, R. Hsu, F. Lee, “Review of Recommender Systems and Their Applications”. Computer Science, 2012.
7. P. Covington, J. Adams, and E. Sargin, "Deep Neural Networks for YouTube Recommendations," Mountain View, California, 2016.
8. A. Bellogín, P. Castells, I. Cantador, “Neighbor Selection and Weighting in User-Based Collaborative Filtering: A Performance Prediction Approach”. ACM Transactions on the Web, No. 12, 2014.
9. S.C. Stephen, H. Xie, and S. Rai, “Measures of Similarity in Memory-Based Collaborative Filtering Recommender System: A Comparison”, 4th Multidisciplinary International Social Networks Conference, 2017.
10. G.S. Milovanovic, “Hybrid content-based and collaborative filtering recommendations with {ordinal} logistic regression (1): Feature engineering”, Data Science Central, 2017.
11. M.G. Vozalis and K.G. Margaritis, "On the Enhancement of Collaborative Filtering by Demographic Data," in Web Intelligence and Agent Systems, Vol. 1, 20063