CSE6012 – IMAGE PROCESSING APPLICATIONS

Digital Assignment – 3 (Final)

Image based furniture recommendation system

HARPAL KAUR DHINDSA – 20MCB0008

School of Computer Science and Engineering, VIT University

Abstract: The goal of this project is to extract furniture items from an image that includes multiple furniture items and then find similar items from the database with the extracted furniture items. The images could be either downloaded from online websites or taken by cameras. The processing process can be divided into three steps: first, certain image segmentation techniques are applied to split out each furniture item; then, an object recognition scheme is applied to each segmented furniture item; finally, a shape descriptor is used to retrieve similar items to the segmented furniture items from a database of each classified category.

Keywords: Image processing, furniture extraction, object recognition, feature matching

Introduction:

In everyday life, individuals may discover some décor plan exceptionally appealing when they are perusing a home enrichment magazine or visiting some all-around planned indoor shops like a bistro, shopping centre, or eateries. Often, they need to discover comparative furniture things and get constant data about them, for example, where they can get them on the web. Therefore, certain methods should be applied to isolate every furniture item from the foundation picture and recover comparative things in the data set. Indoor scene understanding has become a famous exploration point lately because of its importance and challenge in both scholarly examination and mechanical applications. Most related works centre around the errands of scene order, design assessment, and scene parsing. However, scarce works centre around furniture understanding, which is likewise basic for a definitive scene comprehension and recreation.

With the rapid improvement of neural networks and natural language processing these late years, we are now be able to change the customary pursuit standards from text portrayal to visual revelation. A preview of an item recounts a definite story of its appearance, utilization, brand, etc. While a couple spearheading works about picture based pursuit have been applied, the utilization of picture coordinating with utilizing computerized reasoning in the on-line shopping field remains to an extent neglected. Based on this thought, here an attempt to assemble a keen suggestion framework has been made, which takes pictures of items as well as the portrayal text as its info.

Or then again the furniture that they like, where would they be able to get it from or is there a comparable item that is accessible at a lower cost. The Room Décor Assistant uses pictures

which can be from a list or photos taken by the client to discover items that are comparable. It utilizes a mix of image classification and text analysis to give the most fitting ideas. I chose to zero in on Living Room Furniture to begin with, as I felt that furniture in the lounge normally is the primary thing one notices in an individual's home and it something that the vast majority of people are more put resources into when settling on a buy choice opposite different kinds of furniture.

Related Work:

A novel approach of the interactive system, recognized for the customer's product resolve (Shruti R. Pardhi, 2017). A customer needs to capture the image of furniture when they are interested in it, and just upload it on the system to get similar products. With the help of image processing, certain techniques applied to the captured image and system collaboratively learn the customer interest and quickly gives products depending on the customer, which shows not only merchandise variety but also the imaged correspondence of products so that a giant exactness and gratification of product suggestion can be accomplished. A method for object recognition and successive image matching is used, for finding the related furniture items from the database contour-based shape feature matching path is used.

Wang Yong (Wang Yong, IEEE) proposed an approach for the semantic segmentation and structural parsing of modular furniture items, such as cabinets, wardrobes, and bookshelves, into so-called interaction elements. It is challenging to perform such a segmentation into functional units, not just because of the visual closeness of the various components but also on account of their frequent uniformly colored and low-texture appearance. This method addresses these challenges by merging structural and appearance likelihoods of each element and jointly optimizing over shape, relative location, and class labels using Markov Chain Monte Carlo (MCMC) sampling. They proposed a novel idea called rectangle coverings which gives a tight bound on the number of underlying components and henceforth limits the search space.

An approach based on the furniture morphological identification process using cognitive recognition mode of furniture form features (ZHANG Hua, 2014). The research on furniture product image came in recent years, and it focused on how to make the product aesthetic for humans, how to make products more adaptive to the human emotional traits, and emotional elements of designing furniture as well.

Local contour features used are often too generic. They are easily matched to irrelevant parts of an image, while the geometric relationship between different parts of objects is often not quantitatively evaluated. To overcome this issue, Huigang Zhang (Huigang Zhang, 2013) proposed a method for object detection via structural feature selection and a part-based shape model. This method first builds a class-specific codebook of local contour features and then generates feature descriptors by combining context information on object shape. These proposed descriptors are robust against within-class variations and scale changes. Exploring

pairwise image matching via earth mover's distance, discriminative foreground features are selected after an iterative feature weight update process. These selective features from the positive training images are then used to build a part-based shape model. Object detection is performed by matching each testing image against this model.

A method of CBIR, in which features extracted with the help of colour, texture, and shape this approach gives the chance of an effective search. EHG, SIFT, and HOG methods used, with the help of this method of image segmentation gives effective results. As researchers pay more attention to CBIR, in (Leila Kabbai, 2016) unite local and global features with the help of UL-LBP to enhance the strength of image retrieval.

Methodology:

The dataset consists of around 36,000 product images along with their descriptions from overstock.com. Selenium is utilized to web scrape the data from the site. Initially only the URLs are collected for the images and not the actual images. This approach is applied since it is easier to main a list of URLs for manipulating and cleaning data before modelling. The other features for text analysis that the dataset consists are: Description of the image, Style of the furniture, Material of the furniture, Type of class the furniture belongs to, and its colour.

While collecting the data is labelled in batches based on the 5 categories of living room furniture, i.e. Sofas, Ottomans, TV Consoles, Coffee Tables and Armchairs.

The system architecture is shown in the figure below. The flow of the system can be described as follows. The first step is scraping the images from the web, followed by cleaning the data. The data can be divided into two modules the product images, and the text description module. The image data is classified using deep learning method CNN. The test data is analysed using topic modelling method LDA. Both the outcomes are cumulated to get the final result, that is, the recommendations.

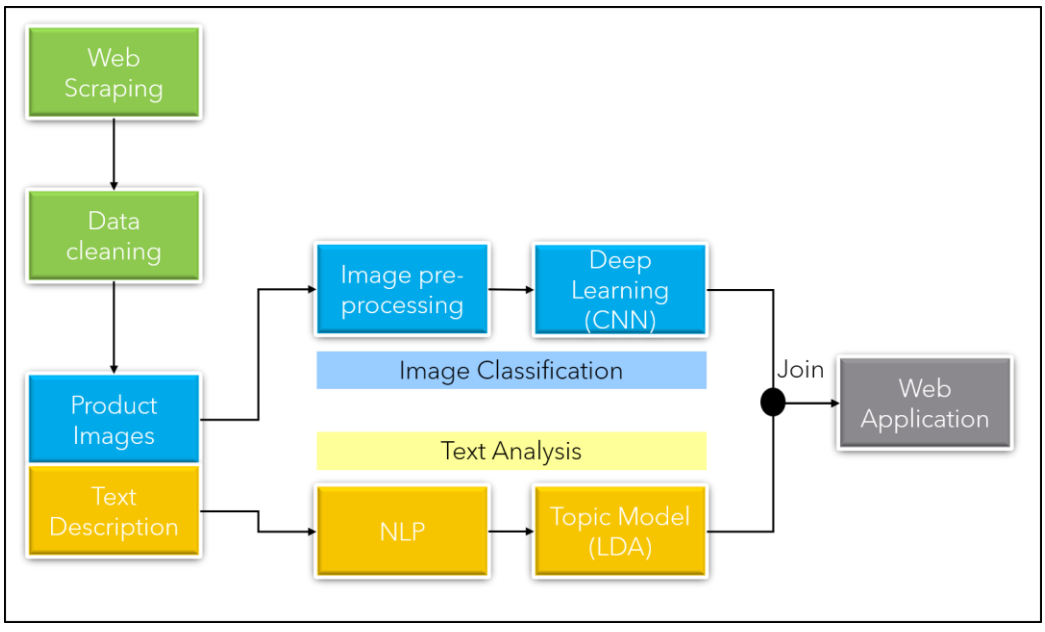

Figure 1. System Architecture

Web Scraping: The data is scraped from overstock.com. First the links are collected from the banners and then used these links to hit individual pages to collect information.

Data Cleaning: The text data is cleaned using pandas, NumPy. However, a few images needed to be removed manually. Images that were zoomed in or had incorrect viewing angles. Although time consuming, after carrying out manual cleaning the model accuracy went up to 91% from 81%.

Splitting the image data: Not all 36,000 images were used to train my Convolutional Neural Network. A balanced set of images from each category/class of 3500 images each were utilized. That is because, 17,500 image that is 3500 per class was enough to train the model. Using more data would have resulted in the model being biased in its learning towards any one type of furniture category since in the real world there is an equal probability that the user picks any category of furniture to find a recommendation.

Deep Learning Convolutional Neural Network: The clean image URLs (17,500) data are used and pre-processed into image matrices (224,224,3) since the VGG16, VGG 19 and ResNet50 CNN models use an input 224 px by 224px across the 3 channels (RGB). After trying out several combinations, the pre-trained VGG16 was finalized after removing its top layers and adding the following layers; 128 Neurons Fully Connected Layer, 64 Neurons Fully Connected Layer, Dropout (0.5) Layer, and Softmax Layer for classifying 5 classes of Living Room Furniture.

Text Analysis: The text data is processed and tokenized the product description with 1 to 3 ngrams with the snowball stemmer. Topic modelling is carried out with LDA across 7 topics. The entire corpus of text descriptions is processed and used to perform log transform of the topic features generated for the recommender.

Making the Recommender

The flow diagram for the recommender is shown in the figure below:

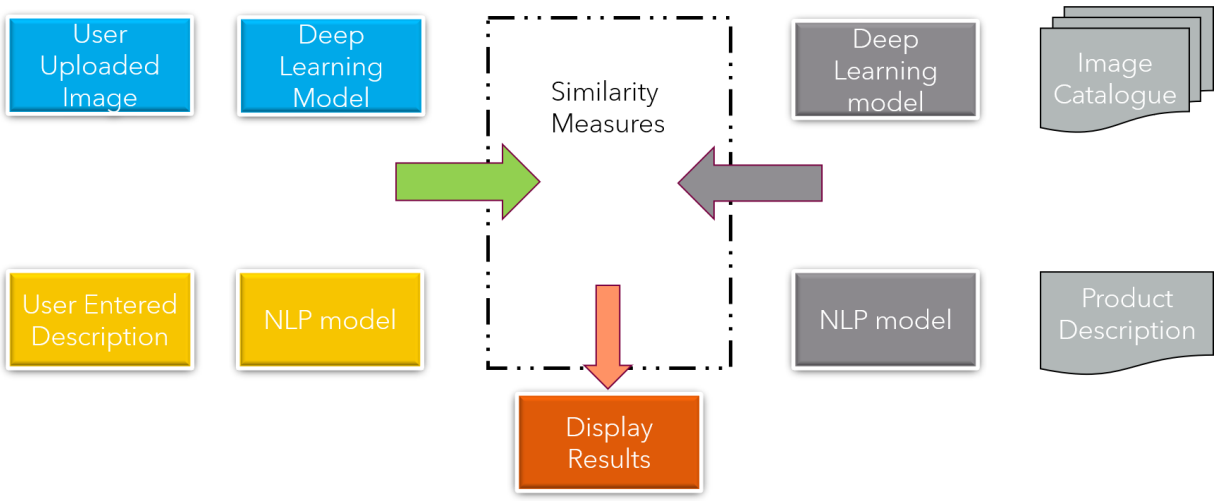

Figure 2. Recommender flow diagram

1. Feature Extraction for all images: After developing the image classification model using Deep Learning CNN, feature extraction is carried out for all the images by using the first Fully Connected Layer of 128 neurons. The model is used to predict these features on the entire image catalogue. Once all the image features are extracted for the entire catalogue this was stored to be later used by the recommender.

2. NLP Features: Features generated during text analysis for the entire data were stored to be used later by the recommender

3. User workflow: The user uploads an image and an optional text description. The user image and text description are processed in the same way the image catalogue and product descriptions were processed to extract features. To find similar products first Euclidean distance is used to find the top 15 closest images and then these are sorted by text similarity using cosine similarity measure. The second stage is omitted if the user does not give a text description. The top 6 closest options are then displayed.

Experiments and Results:

On running the application on flask, the home page is shown in Figure below.

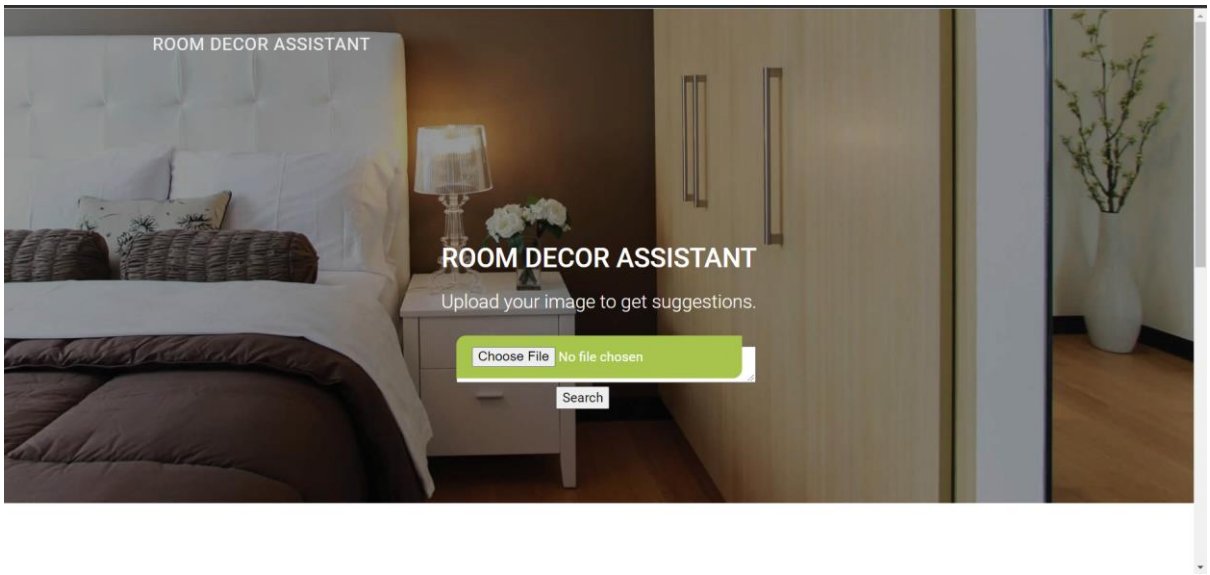

Figure 3. Home Page

The layout of the resultant recommendation images is shown in Figure 4..

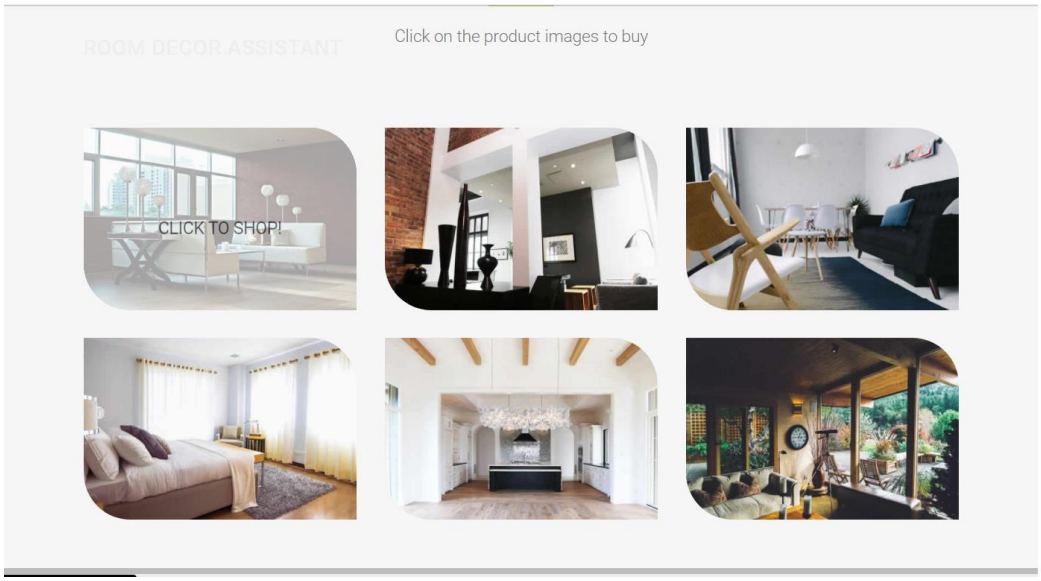

Figure 4. Home Screen

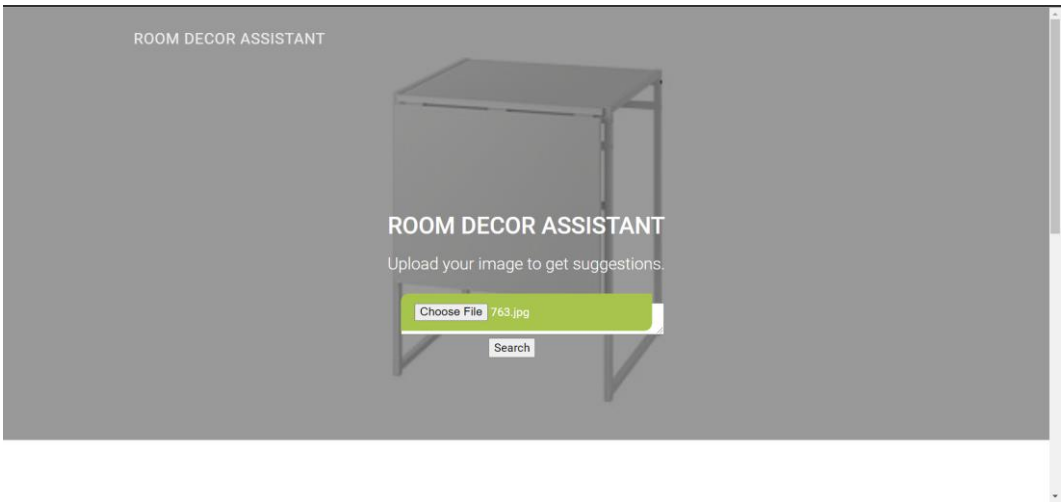

Figure 5. Selecting an image of a table

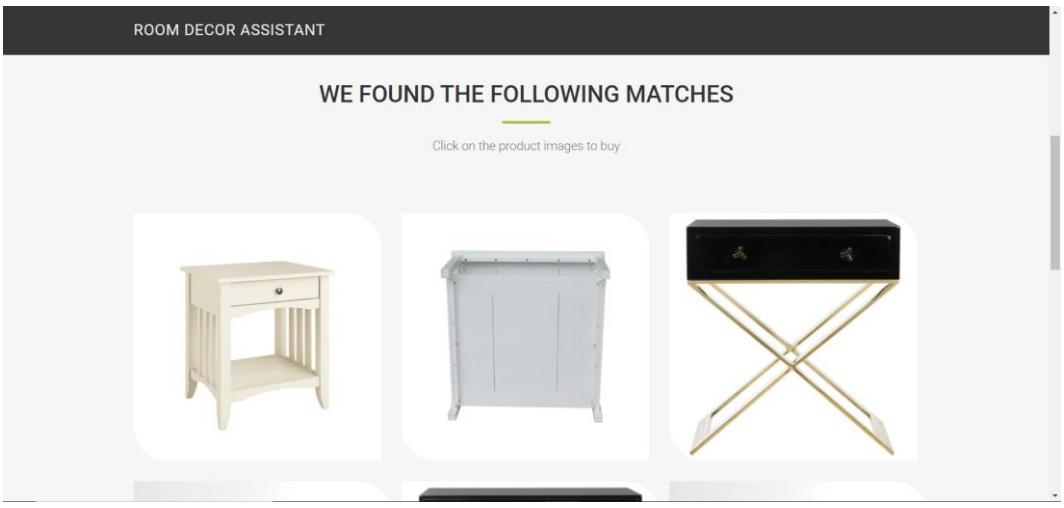

Figure 6. Recommended tables

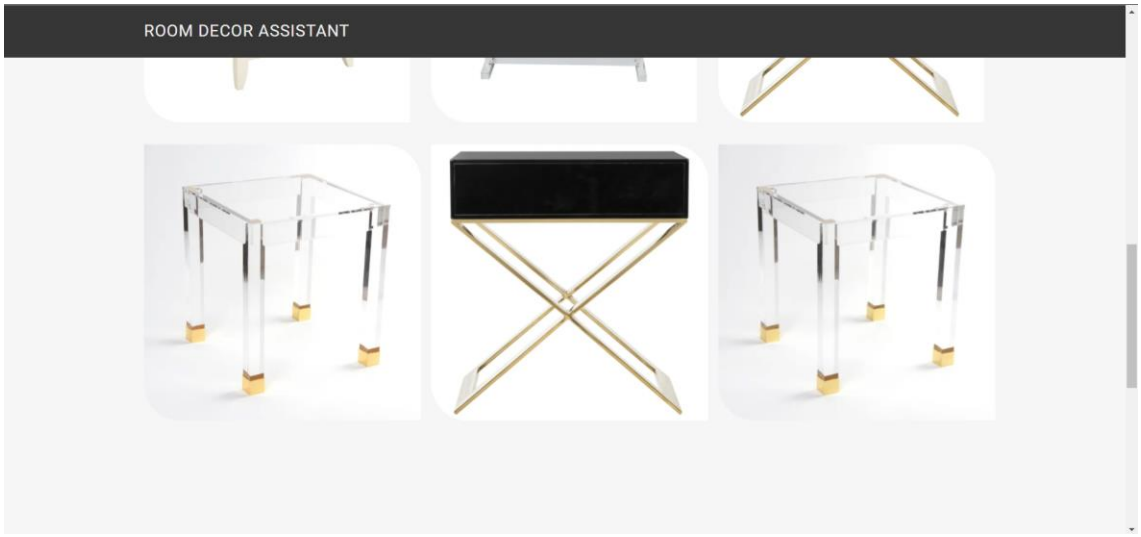

Figure 7. Recommended Images

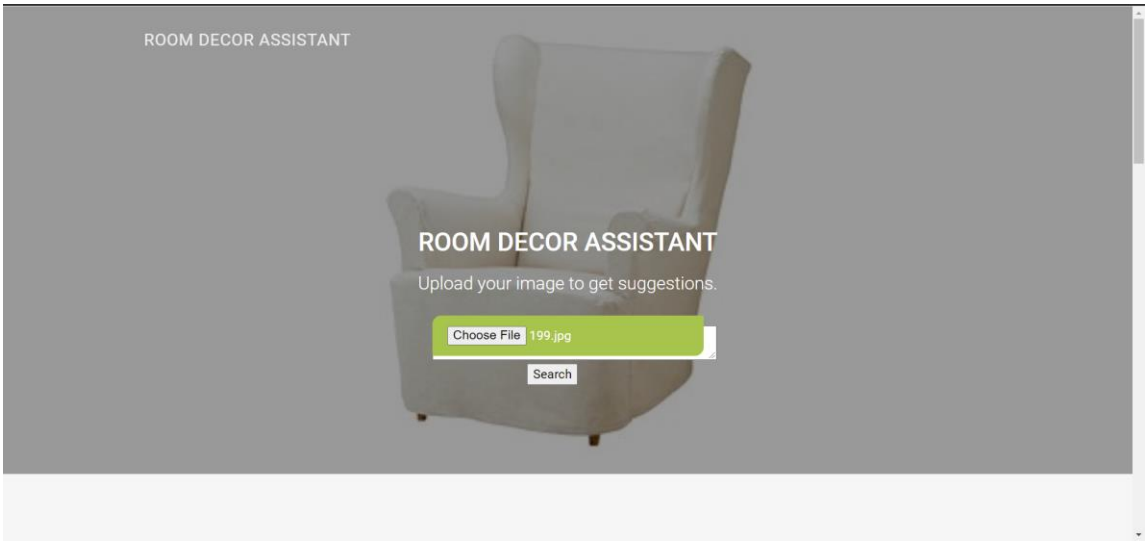

Figure 8. Selecting an image of a chair

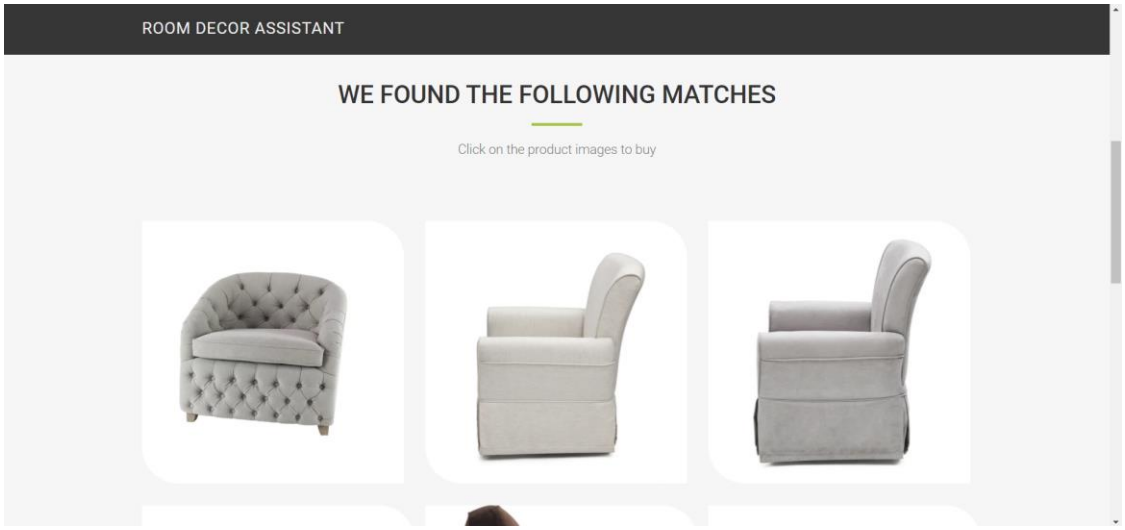

Figure 9. Recommended images

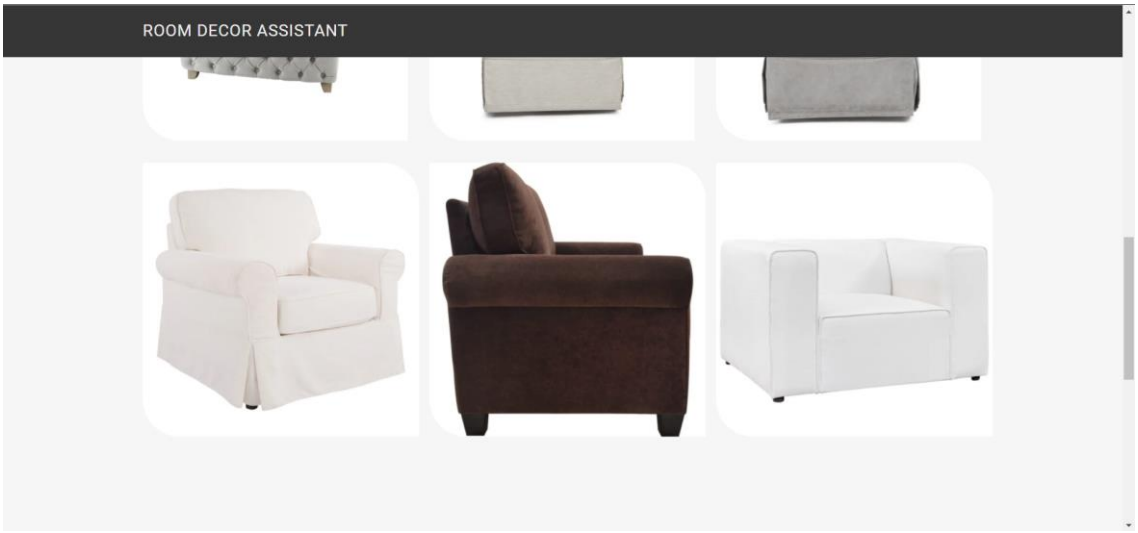

Figure 10. Recommended Images

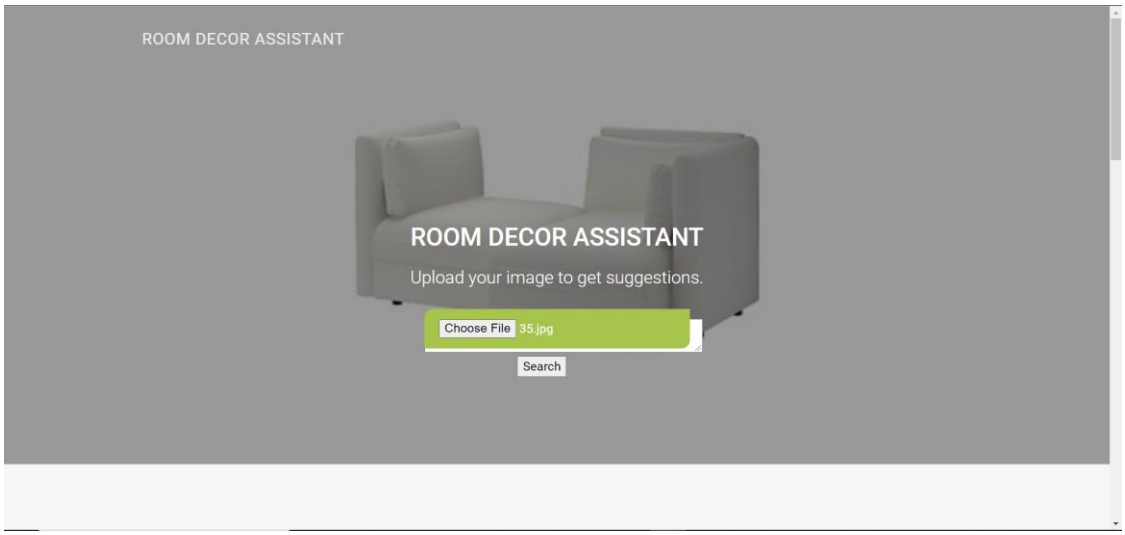

Figure 11. Selecting the iamge of a sofa

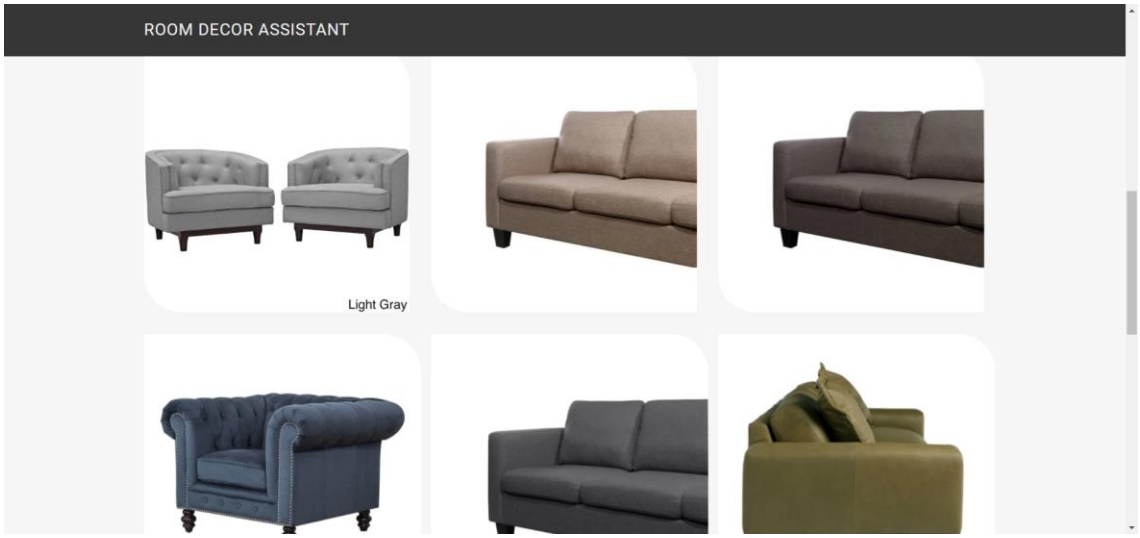

Figure 12. Recommended Images

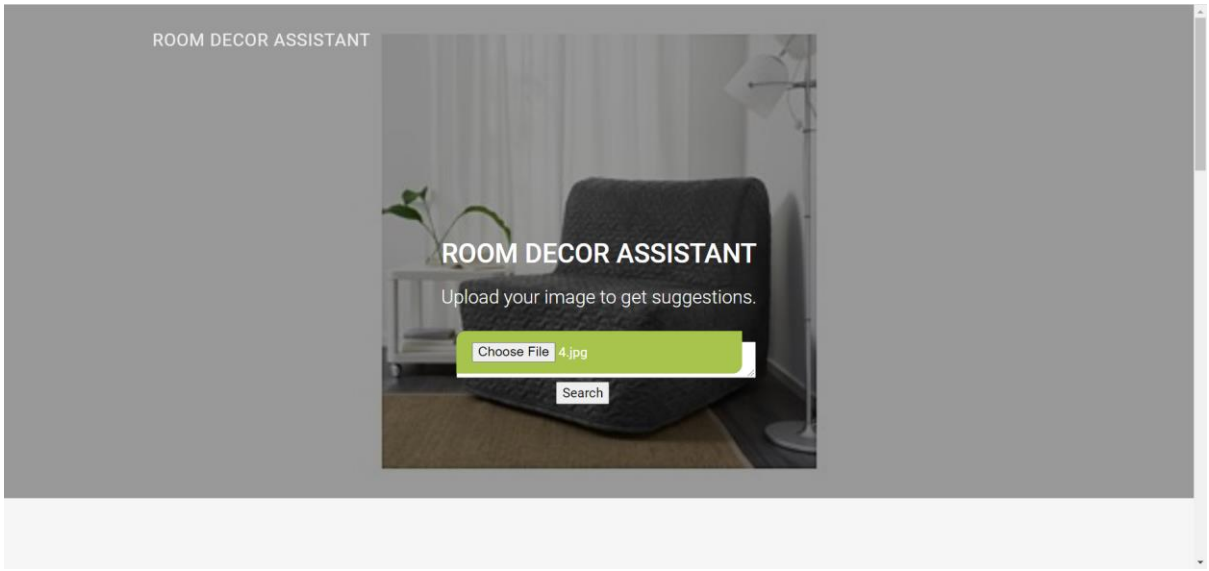

Figure 13. Selecting the iamge of a sofa

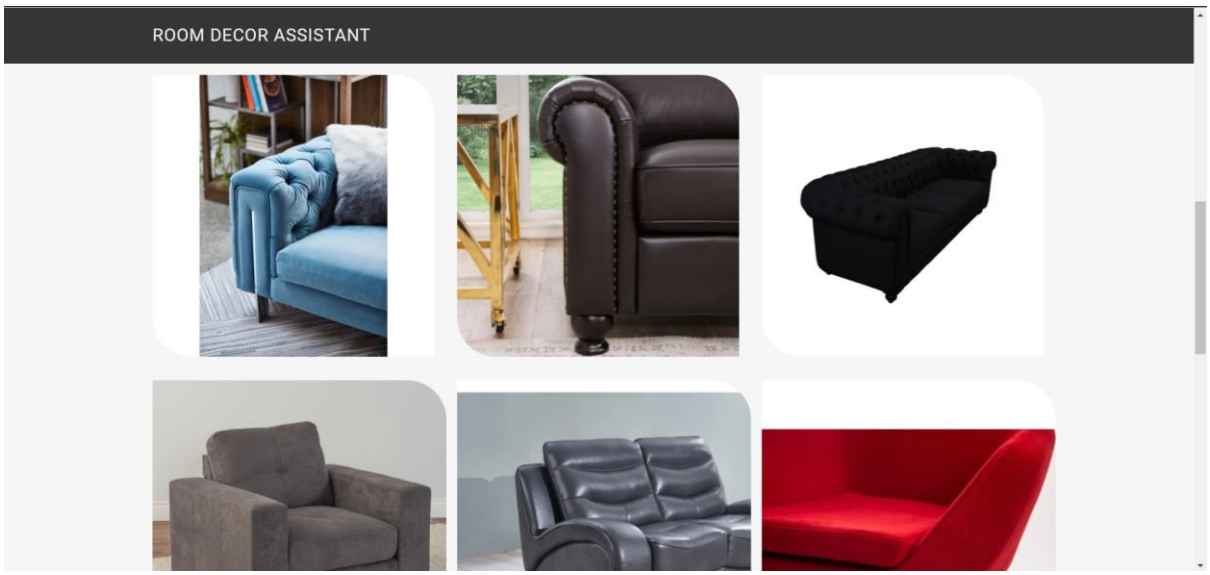

Figure 14. Recommended Images

Text analysis, topic modelling output:

The figure 13 shows the 6 topics, each with their feature names.

```
------PERPLEXITY------
96.44484755778346
```

```
no_top_words = 10
tf_feature_names = tf_vectorizer.get_feature_names()
display_topics(lda_tf, tf_feature_names, no_top_words)
```

```
Topic: 0
centuri , mid , mid centuri , centuri modern , mid centuri modern , modern modern , modern modern contemporari , centuri modern modern , leg
, grey
Topic: 1
sofa , contemporari , comfort , seat , grey , live , modern contemporari , cushion , leather , room
Topic: 2
tabl , cocktail , cocktail tabl , tabl coffe , tabl coffe tabl , coffe , coffe tabl , accent , accent tabl , cocktail tabl coffe
Topic: 3
tabl , coffe , coffe tabl , wood , brown , rustic , home , room , end , live
Topic: 4
ottoman , chair , pouf , armchair , contemporari , home , seat , storag , blue , comfort
Topic: 5
stand , consol , entertain , storag , media , contemporari , brown , shelv , modern contemporari , inch
Topic: 6
tabl , coffe , coffe tabl , contemporari , modern contemporari , contemporari coffe , contemporari coffe tabl , glass , modern contemporari
coffe , black
```

Figure 13. Topic Modeling output

Figure 15-17 are the pyLDAvis visualizations. It is designed to help users understand the topics in a topic model that has been fit to a corpus of text data. The package extracts data from a fitted LDA topic model to update an interactive web-based visualization.

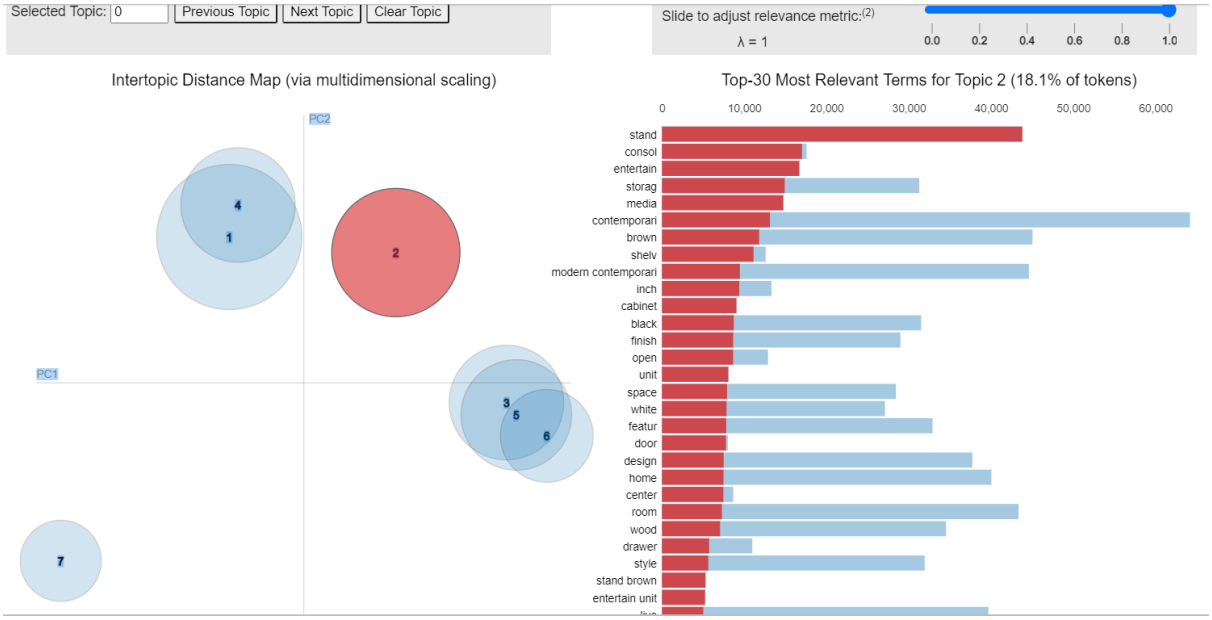

Figure 15. Topic 2 frequency visualization

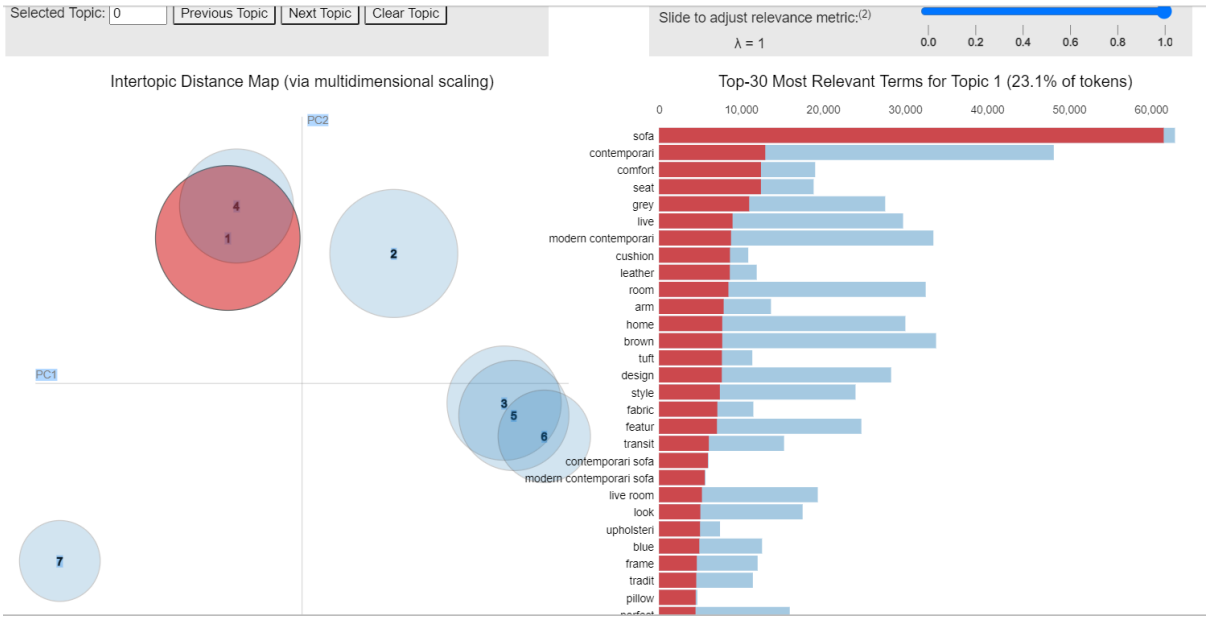

Figure 16. Topic 1 frequency visualization

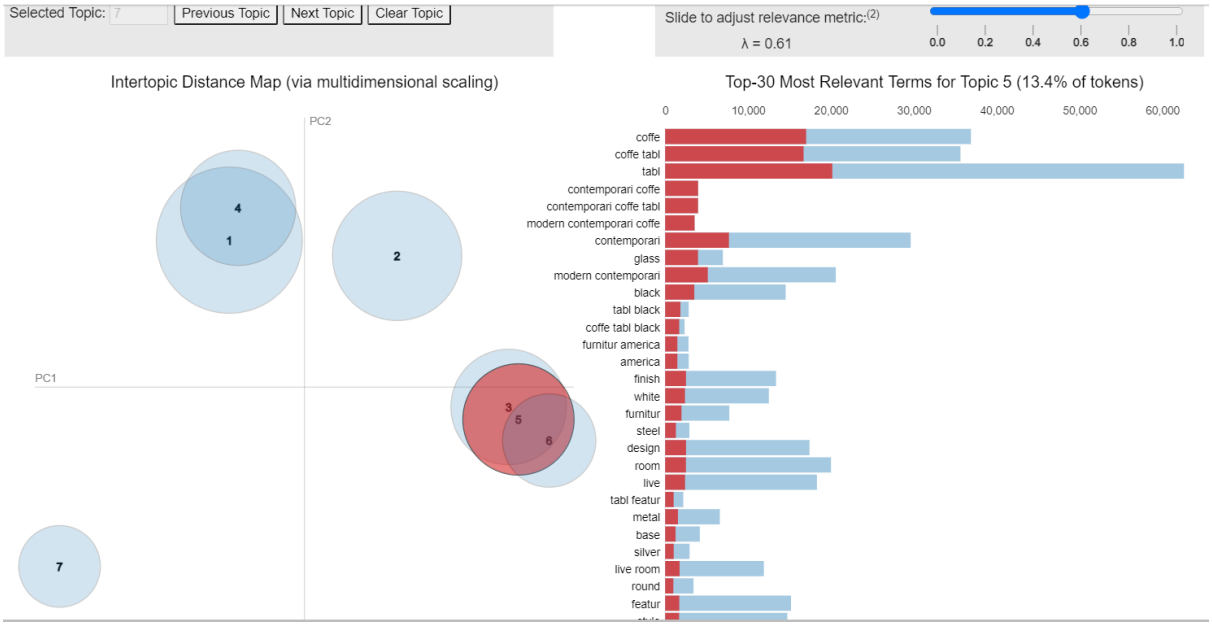

Figure 17. Topic 3 frequency visualization

Conclusion:

The results show that text analysis is not as effective in this scenario. The reason being that a lot of descriptions were very subjective and sales oriented with words like elegant, beautiful, etc. A low max_df of 70% was used to get rid of words that kept occurring very frequently to extract meaningful words. Viewing angles, background and lighting conditions can affect model outcomes and going forward this is a challenge that needs to be addressed more fully. More tolerance for bad photography, learning from user images, adding more categories and products from other websites, and add hard filters to avoid bad suggestions. A multilevel

(nested) recommendation can be added where user can select the images he or she likes and the application can utilize these to find further recommendations.

References:

[1] Sui Ting Ting, Z. K.-Y. (2019). "Interest Points guided Convolution Neural Network for Furniture Styles Classification". *The 2019 6th International Conference on Systems and Informatics (ICSAI)*.

[2] Luisa F. Polan´ıa, M. F. (2020). "Learning Furniture Compatibility with Graph Neural Networks". *IEEE*.

[3] Shruti R. Pardhi, K. W. (2017). "Extraction and retrieval of furniture from designing decoration and furniture database". *ICCCI*.

[4] Leila Kabbai, M. A. (2016). "Content Based Image Retrieval using Local and Global features descriptor". *IEEE*.

[5] Luyang Chen, F. Y. (2016). "Image-based Product Recommendation System with Convolutional Neural Networks" . *IEEE*.

[6] Tobias Pohlen, I. B. (2016). "Semantic Segmentation of Modular Furniture". *IEEE*.

[7] Wang Yong, F. D. (IEEE). "Design and implementation of real-time image monitoring system based on Android mobile terminal". *2014*.

[8] ZHANG Hua, S. H. (2014). "Study of Pattern Recognition of Morphological Characters for Furniture Image Cognition". *IEEE*.

[9] Huigang Zhang, X. B. (2013). "Object Detection via Structural Feature Selection and Shape Model". *IEEE*.

[10] Tang, J. (2010). "A Color Image Segmentation algorithm Based on Region Growing". *IEEE*.