Marathwada Mitra Mandal's College of Engineering, Pune

Image-based Product Recommendation System for E-Commerce Websites

Bhavika Mauli Karale

Under the guidance of
Prof. Pooja Dhule

Roll No: TC178

Date: 15/06/2021

Index

Introduction

- In the online marketplace world, companies make use of user history and product data to make recommendations to user.
- The data collected to make these recommendations are search terms, purchase history, product category, items in cart information and many more.
- Most commonly, recommendation engines make use of text based keyword matches to identify products that consumers may like.
- Another alternative and less widely used technique is to make use of images to compare product similarity. Images may be better representative of consumers interest than text.

Objective

- To make an interactive setup that is convenient and reliable for querying related products.
- Implement an idea of searching for products in an online shopping system using an image-based approach.
- A user can provide, select, or click an image, and similar image-based products will be presented to the user.

Related works

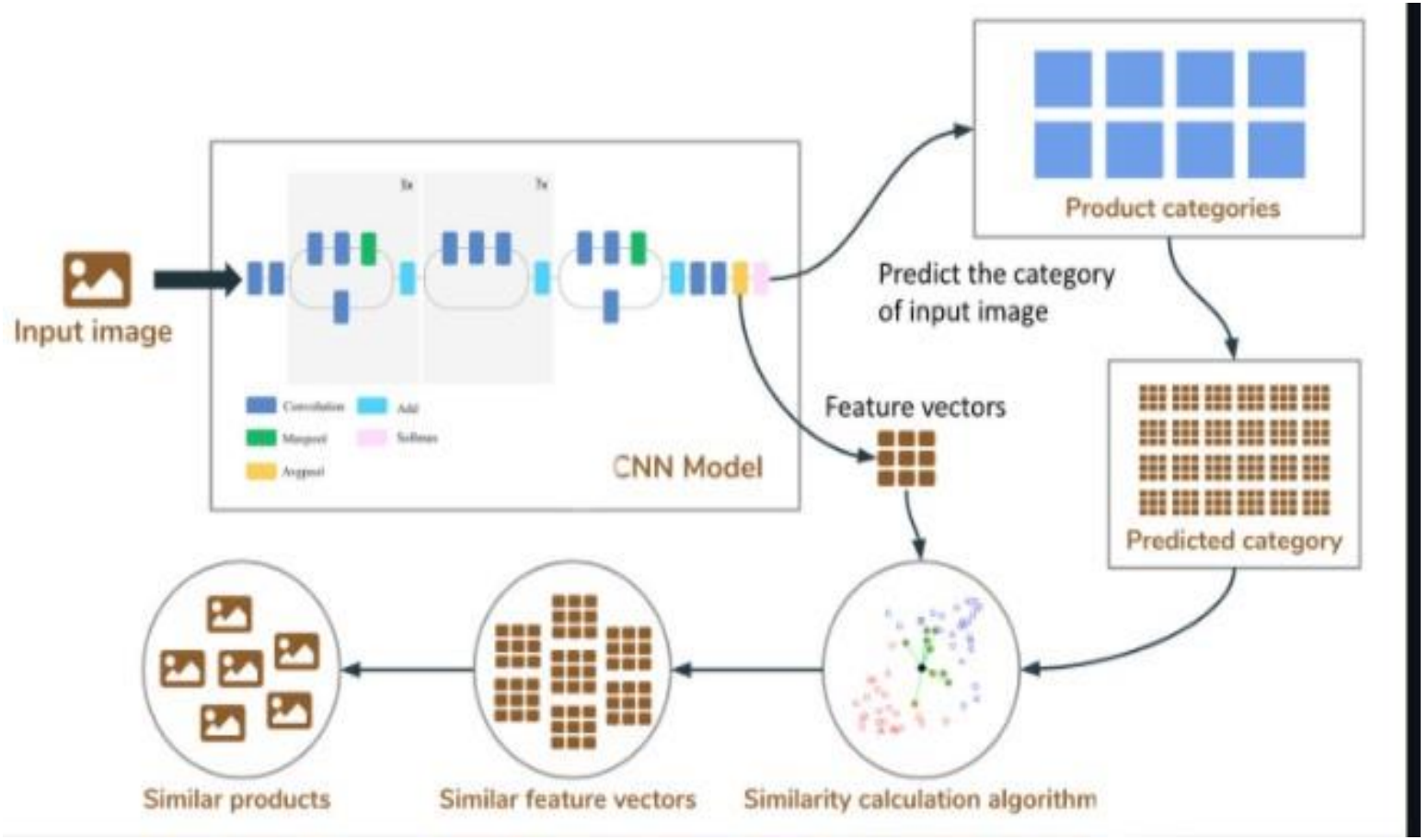
Title	Authors	Concept	Advantages	Disadvantages
Image based service recommendation system: A JPEG coefficient RFS approach	Farhan Ullah, Bofeng Zhang, Rehan Ullah Khan	Proposed model uses a Random Forest classifier in the classification phase and JPEG coefficient vectors are used to calculate similarity scores by means of Euclidean distance and Cosine similarity in the recommendation phase to provide most closely matching images to the user input.	For product's class learning, the RF classifer is used. For feature extraction from images, the JPEG coefficients are used as image features	For performance enhancements, the RF model needs to be integrated into the DL setup

Image based product recommendation system with convolutional neural networks	Luyang Chen, Fan Yang, Hequing Yan	Authors have discussed how deep convolutional networks like AlexNet and VGG models with a baseline SVM model can be used to solve the classification problem and Jaccard similarity is used to solve the recommendation problem.	A CNN makes predictions by looking at an image and then checking to see if certain components are present in that image or not. If they are, then it classifies that image accordingly	They completely lose all their internal data about the pose and the orientation of the object and they route all the information to the same neurons that may not be able to deal with this kind of information
Image Based Fashion Product Recommendation with Deep Learning	Hessel Tuinhof, Clemens Pirker and Markus Haltmeier	In the proposed model, the authors train a convolutional neural network (CNN) which is used as a problem-specific feature extractor, where the features serve as inputs for the ranking system. The knearest algorithm (k-NN) is used for ranking in feature space.	A neural network classifier is used as a data-driven, visually-aware feature extractor. Initialization strategies using transfer learning from larger product databases are presented	Accuracy based on a small dataset. Problem of cold start not addressed

Image Based Search Engine Using Deep Learning.	Surbhi Jain and Joydip Dhar	The paper proposes an architecture of Deep Learning for CBIR systems. A pretrained CNN model, that is, Inception-v3 model, a GoogleNet deep architecture is applied on the dataset for classification. Further, Euclidean Distance is used as similarity metric and closely matching products are retrieved.	Use of transfer learning. Extracting the last-but-one fully connected layer from the retraining of GoogleNet CNN model served as the feature vectors for each image, computing Euclidean distances between these feature vectors and that of our query image to return the closest matches in the dataset.	The paper did not consider the problem of sparse data. The approachesconsidered would have lot of sparse data and did not mention how they would deal withsuch redundant sparse matrix

'येथे बहुतांचे हित ।'

Proposed System Architecture



Algorithm

Phase 1: Algorithm for Image Classication:

1. Create a function to download the image links from the dataset.
2. Use the above function here to store all the images in the dataset into array.
3. Convert the response variable into numbers.
4. Split the data into training and testing sets.
5. Normalize pixel values to be between 0 and 1.
6. Set the seed and add the convolutional layers.

7. Add the atten and last dropout layers.
8. Compile the model.
9. Fit the model.
10. Perform prediction on the test image.
11. Display confusion matrix for results.
12. Convert the test labels into a list.
13. Create a function which picks random images and identies the class to which the image belongs.

Phase 2: Algorithm for Image Recommendation:

1.Load the model.

2.Remove the last layers in order to get features instead of predictions.

3.Print the layers of the CNN.

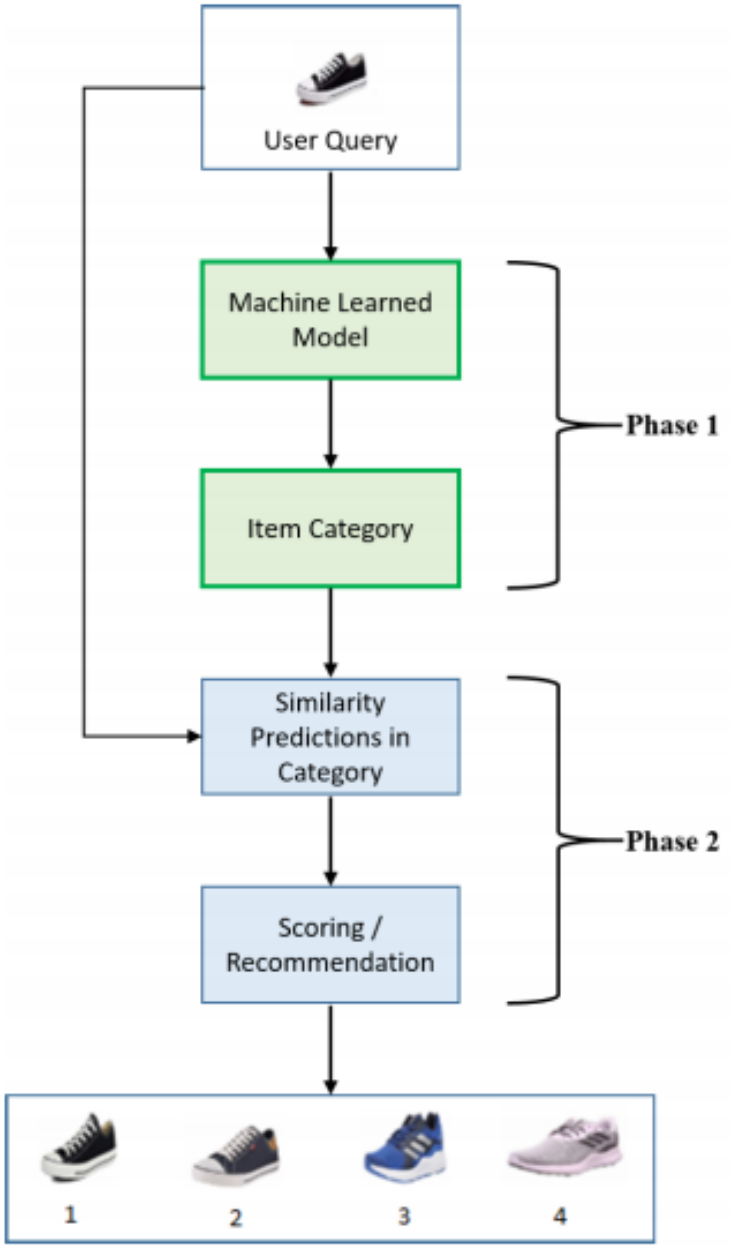
4.Load the images that were downloaded.

5.Use regex to extract only the pids from le names.

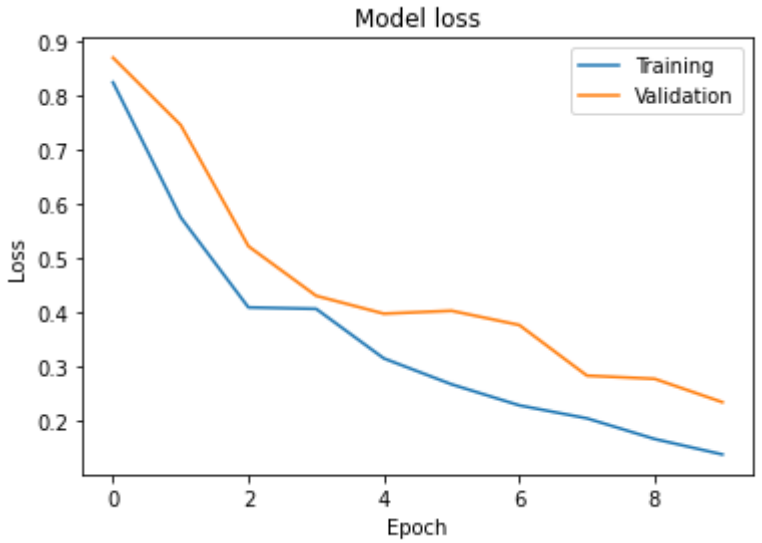
6.Add the pid and the image URL to a dataframe.

7. Load all the images and prepare them for feeding into the CNN.
8. Set the image size to 224*224.
9. Convert the images to array.
10. Extract the images features.
11. Compute cosine similarities between images.
12. Store the results into a pandas dataframe.
13. Create a function to retrieve the most similar products for the given one.

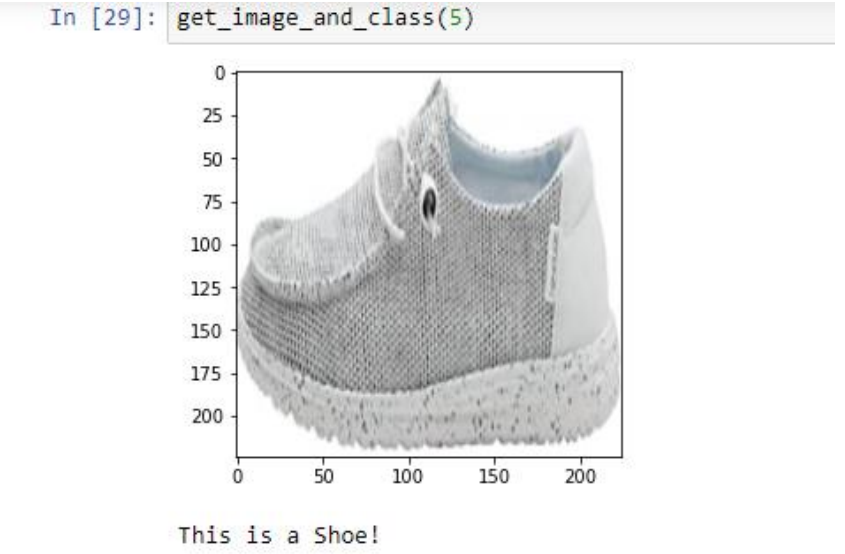
Generic Flow of proposed approach



Results(Feasibility)



CNN Model accuracy of 89.9% was achieved



Classification Result

```
In [36]: image_recommend_1(500,5)
```

```
--------------------------------------------------------------------------
Original product:
--------------------------------------------------------------------------

Product ID :  500
```


```
In [36]: image_recommend_1(500,5)
```

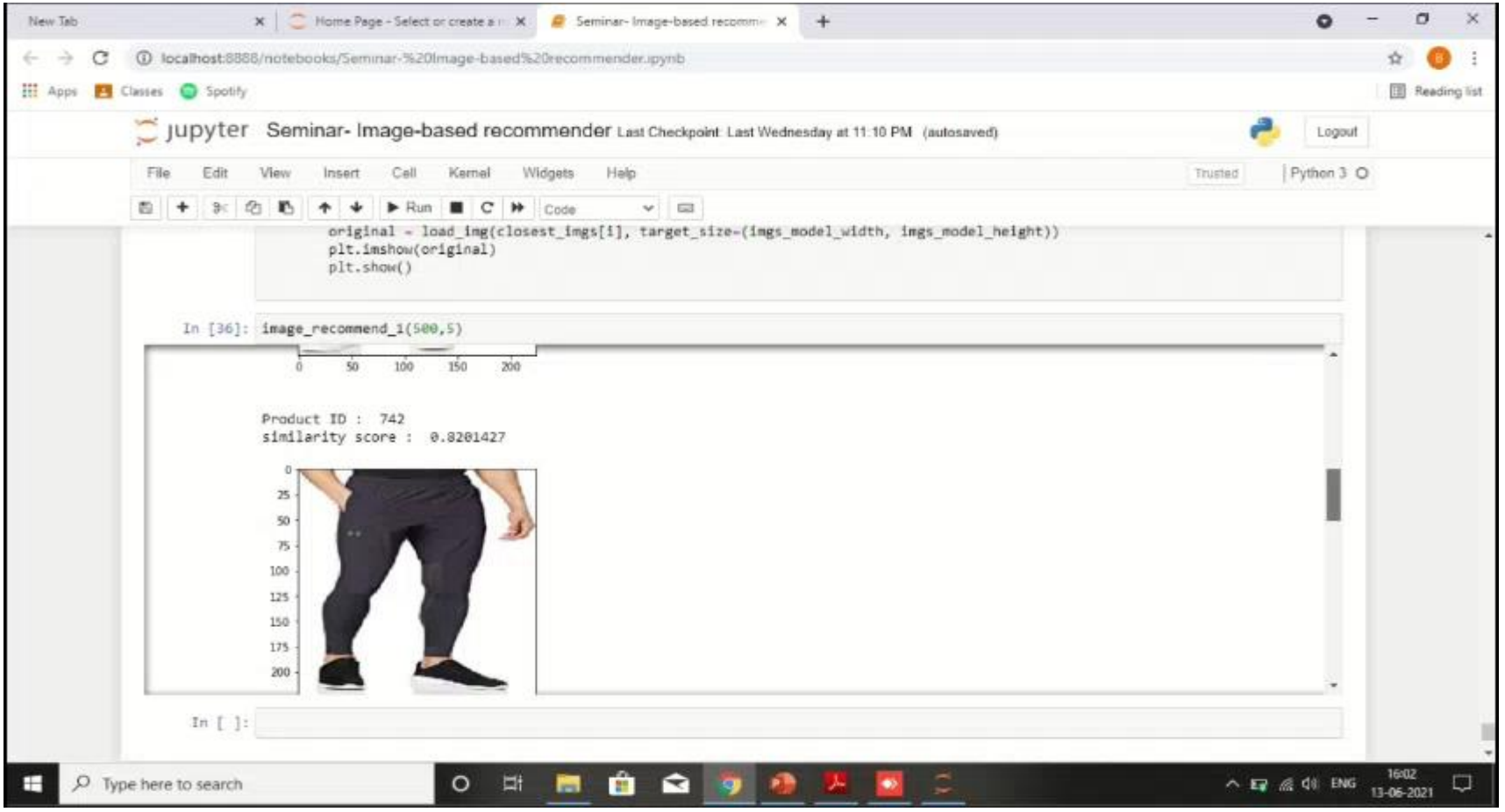
```
--------------------------------------------------------------------------
Most similar products:
--------------------------------------------------------------------------

Product ID :  501
similarity score :  1.0000001
```

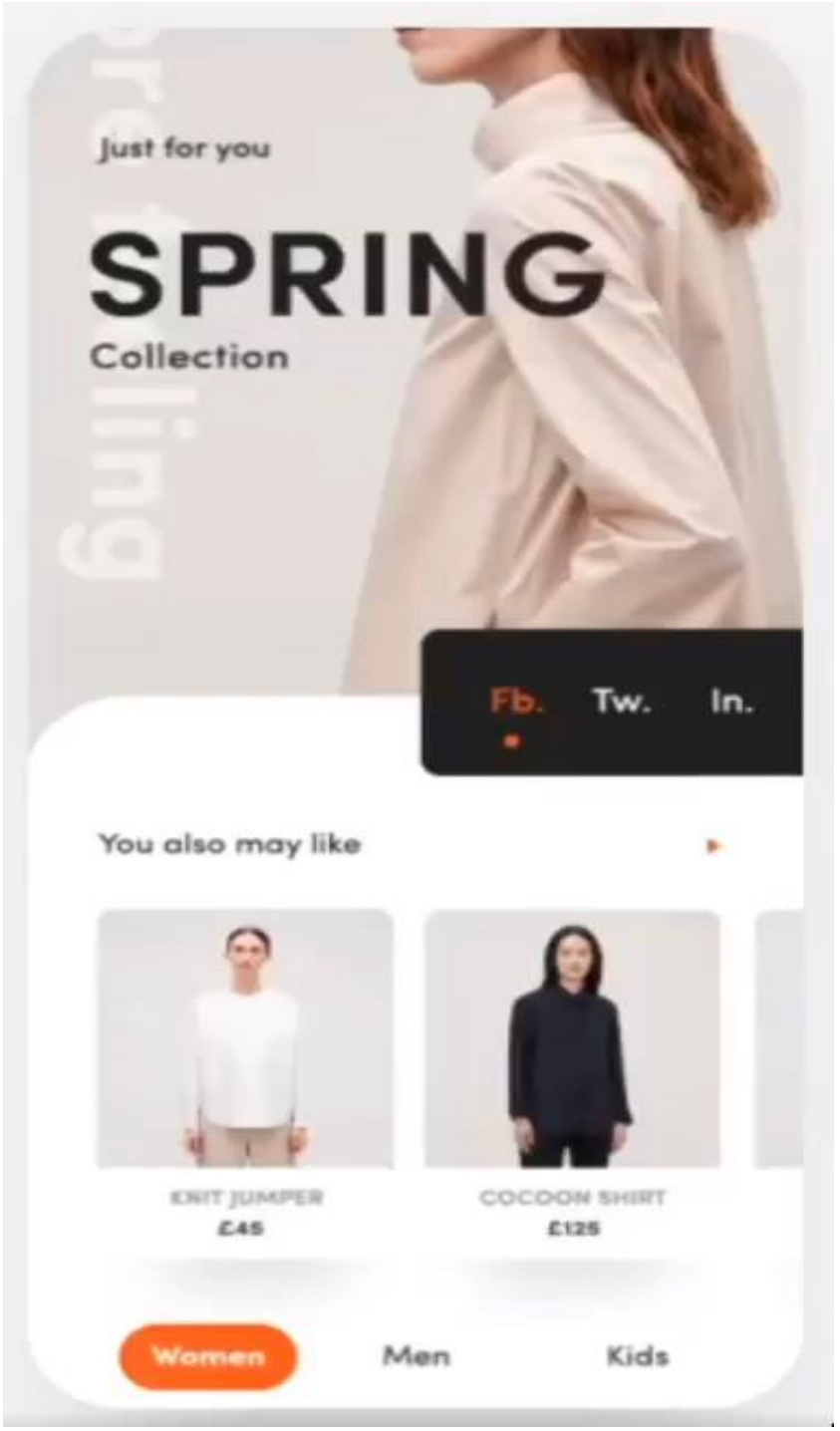
RECOMMENDATION RESULT

As depicted in the above screenshots, 5 similar products are recommended based on the original/input image.

'येथे बहुतांचे हित ।'

Demo
(Screenshot/video of product)




just for you
SPRING
Collection
Fb.
Tw.
In.
You also may like
KNIT JUMPER
£45
COCOON SHIRT
£125
Women
Men
Kids

App Demo

Applications(usability)

- Visual Search for Improved Product Discoverability
- Image Classification for Websites with Large Visual Databases
- Image and Face Recognition on Social Networks
- Interactive Marketing and Creative Campaigns
- Automated Image Organization – from Cloud Apps to Telecoms

Conclusion

- A two-phase recommendation model is discussed, which performs classification and recommendation in respective phases.
- Convulational Neural Network is used for classication purpose and similarity scoring is done using Cosine Similarity for recommendation phase. In some cases, the machine learning model is integrated into a Deep learning setup for enhanced results.
- Hence, better recommendations can be provided to the user based on visual search.

Future work

- Build a recommendation engine that incorporates both the image-based and text-based methods.
- Try to train the model on a larger amount of data using batches. This can potentially increase the accuracy of the model
- Create a web app to deploy the model- Classification & Recommendation

References

- Farhan Ullah, Bofeng Zhang, and Rehan Ullah Khan, "Image-Based Service Recommendation System: A JPEG-Coefficient RFs Approach",IEEE, 2020.
- Luyang Chen, Fan Yang and Heqing Yang, "Image-based Product Recommendation System with Convolutional Neural Networks.", Stanford University cs213n reports, 2017.
- Hessel Tuinhof, Clemens Pirker and Markus Haltmeier, "Image Based Fashion Product Recommendation with Deep Learning ", July 2018.
- Surbhi Jain and Joydip Dhar, "Image Based Search Engine Using Deep Learning.", Proceedings of 2017 Tenth International Conference on Contemporary Computing (IC3), 10-12 August 2017.