Assumption:
Image and review are the parts of the given data set

Image Feature Extraction

Approach:
-store .csv into a data frame using pandas library
-Download each image.
- Apply pre-processing techniques such as altering contrast, resizing, geometrical orientation, random flips, brightness, and exposure on each image.
-Use a pre-trained Convolutional Neural Network Architecture VGG16 to extract relevant features from the images.
-we get a vector that contains the feature values of each image.
-we normalize the feature values

Result:
we get a vector that contains the feature values of each image.
Assumptions:
-Some images have not got downloaded. row number 67,110,523,701,860,936

Text Feature Extraction:

Approach:
-store .csv into a data frame using pandas library
-Apply Lower-Casing, Tokenization, removing punctuations, Stop Word Removal, Stemming and Lemmatization on the given text reviews in the data.
-Calculate the Term Frequency-Inverse Document Frequency (TF-IDF) scores for the textual reviews.

Methodologies:
store .csv into a data frame using the pandas library
For each review convert the text into lowercase
Remove punctuation, in this replace with ‘ ‘
Perform tokenization, by using nltk using word_tokenize(text)
Remove stop words using nltk library
Removal, Stemming, and Lemmatization on the given text reviews
We calculate term frequency for each review. For each term, we divide the frequency of occurrence of that term in a particular review by the number of word tokens in that review.
For inverse document frequency we have to do log(N/n), where N is the number of reviews and n is the number of documents containing that word.
After that, we calculate the TF-IDF score for each review and make a vector of TF-IDF for each review.
Result:
We get the TF-IDF score vector for each review

Assumptions:
We utilize libraries such as NLTK and BeautifulSoup for data preprocessing.
We removed numbers and special characters.
We are taking 1000 as N.

Image Retrieval

Approach:
-For the given input of the URL of the image and its review, find the corresponding feature vector and then compare its vector with other image vectors by finding the cosine similarity score.
-Get the top three images whose similarity scores are maximum
-find the cosine similarity of the given input review with the top three reviews of the images retrieved.

Methodologies:
input, URL of the image, and its review.
Find the image feature vector by searching into the extracted feature vector that we have calculated using image feature extraction. And the same for input review.
Now we go through each image feature vector and find the cosine similarity with the image feature vector of the input image and store them in list.
Sort the list by descending order and take the top three values and their corresponding images.
Return image URLs and their corresponding reviews, along with their cosine similarity score of image and text(by finding the cosine similarity score of the input review and top three image reviews)

Result:
URLs of images
Their corresponding reviews
The cosine similarity score of images
The cosine similarity score of reviews

Assumption:
Did not rank images based on combined similarity scores but based on the cosine similarity score of images.
Although calculated composite similarity by taking the average of cosine similarity scores of image and text.

Text Retrieval:
Approach:
-For the given input of the URL of the image and its review, find the corresponding tf-idf vector and then compare its vector with other review vectors by finding the cosine similarity score of the text.
-Get the top three reviews whose similarity scores are maximum

-find the cosine similarity of the given input image with the top three images of the reviews retrieved. In case multiple images are found corresponding to the review then we take the average of the cosine similarity score of the image.
Methodologies:

input, URL of the image, and its review.
Find the tf-idf vector of input review by searching into the tf-idf vector that we have calculated using Text feature extraction. If can not find then calculate again.And the same for input image.
Now we go through each tf-idf vector and find the cosine similarity with the tf-idf vector of the input review and store them in a list.
Sort the list by descending order and take the top three values and their corresponding reviews.
Return a list of image URLs and their corresponding reviews, along with their cosine similarity score of the image(average in case of multiple images found corresponding to a review) and text cosine similarity score.

Result:
List of URLs of images
Their corresponding reviews
The cosine similarity score of images
The cosine similarity score of reviews

Assumption:
Did not rank the review based on combined similarity scores but based on the cosine similarity score of reviews.
Although calculated composite similarity by taking the average of cosine similarity scores of image and text.

Combined Image and Text Retrieval:

Approach and assumption:

For the given input of the URL of the image and its review, find the corresponding tf-idf vector and get image feature vector then find the composite similarity score of each image and review and store in a list.
-Get the top three (I,R) pairs whose composite scores are maximum
-return the cosine similarity of the given input image with the top three chosen (I,R) pairs retrieved. In case multiple images are found then we take the average of the cosine similarity score of the image.
-return cosine similarity score of review
-Ans composite similarity score

Image retrieval is good because it gives better combined score as compare to text retreival

Improvement:

We can improve our text processing by using methods like BM25 mode
We can also improve image processing by using other methods in pre-processing

Challenges

We always do not get similar images due to some constraints of the pre-trained CNN model.
Sometimes we do not get the desired output due to the features it has retrieved.
The distribution of your target images might differ from the pre-trained dataset.

It will not work properly if the given image is not part of the dataset