Report

Image Feature Extraction

Firstly, the data was read using the CSV file, and URLs were accessed and tried to retrieve the data. If not accessible, an error was printed.
The extracted image was then pre-processed by resizing, flipping, and enhancing contrast and brightness.
Then, features were extracted using the ResNet50 model, which will help calculate cosine similarity.

Relevant mapping dictionaries were made to access data in future
All the relevant data has been dumped in pickle files.

The following URL's were not able to be accessed:
['https://images-na.ssl-images-amazon.com/images/I/71F3npeHUDL._SY88.jpg',
'https://images-na.ssl-images-amazon.com/images/I/71wHUWncMGL._SY88.jpg',
'https://images-na.ssl-images-amazon.com/images/I/71B8OOE5N8L._SY88.jpg',
'https://images-na.ssl-images-amazon.com/images/I/81SX3oAWbNL._SY88.jpg',
'https://images-na.ssl-images-amazon.com/images/I/718niQ1GEwL._SY88.jpg',
'https://images-na.ssl-images-amazon.com/images/I/61OboZT-kcL._SY88.jpg',
'https://images-na.ssl-images-amazon.com/images/I/710a2Pyh5lL._SY88.jpg',
'https://images-na.ssl-images-amazon.com/images/I/816NMd0LexL._SY88.jpg']

The following review-ids were hence deleted/discarded:
2235, 3317, 2912, 2265, 2088, 3474

Text Feature Extraction

The reviews were extracted and pre-proccessed with the steps involving tokenization, stop-words and punctuation removal, stemming, lemmatization.
Pre-processed data is then used to calculate the tf-idf values for and dumped in pickle files for easier and faster access.

Image and Text Retrieval

A couple of 2D matrices are made for storing the cosine similarity of image across different review ids and for text reviews with each other.
These are pre-computed and stored

Once the user query comes in, we get the relevant review-id from the database, and access it's row from the matrix, and get the top 3 matching products based on cosine-similarity score.

*Note- For images with products having several corresponding images average across all has been taken

Results and Analysis

I tried several different test cases, and in all of them, 'Image Retrieval' results were better than 'Text Retrieval', as seen from *'Composite Similarity Score'.*

A possible reason is that the reviews for a similar product can be from a vast range and can be described in several different manners to mean the same thing, but when taking features from images, similar products will look similar irrespective of user, just the background can differ to some extent.

Critical challenges include the time taken to process the images and compute the cosine similarity values.
A small improvement was to compute the cosine similarity value for all the products beforehand to ease and quicken user queries.
Another method can include reducing the number of comparisons i.e., instead of comparing to all products using a heap like data structure to reduce the number of comparisons