CSE508 Information Retrieval Winter 2024 Assignment-2 Report

AKASH KUMAR
MT23012

Date of Submission:-14-03-2024

Table of Contents

1. Introduction

In this Assignment, we design and implement a Multimodal Retrieval System that utilises both text and image inputs to retrieve relevant information. The system is aimed at efficiently retrieving similar images and reviews based on a given query, leveraging techniques such as image feature extraction, text feature extraction, and similarity calculations.

2. problem statement

The assignment requires the development of a Multimodal Retrieval System using a dataset consisting of image URLs and corresponding text reviews for a given product. The system is expected to perform image and text feature extraction, calculate similarity scores, and retrieve relevant information based on a provided query.

3. Approach

3.1. Data Preprocessing

For text

- HTML tag removal using BeautifulSoup.
- Lowercasing the text.
- Tokenization using NLTK's word_tokenize function.
- Removing stopwords and non-alphabetic tokens.
- Lemmatization using NLTK's WordNetLemmatizer.
- Calculating Term Frequency (TF) for each word in the text.
- Calculating Inverse Document Frequency (IDF) for the entire dataset.
- Calculating TF-IDF scores for each review.

Image Preprocessing:

- Downloading images from URLs using requests.
- Resizing images to the input size required by VGG16 (224x224).
- Converting images to RGB format and normalizing pixel values.
- Extracting features from images using a pre-trained VGG16 model.

- Flattening the feature vectors.

3.2. Text Feature Extraction and TF-IDF Calculation

Text Feature Extraction:

- Preprocess the text by removing HTML tags, converting to lowercase, tokenizing, removing stopwords, and lemmatizing the words.
- Calculate the Term Frequency (TF) for each word in the preprocessed text.
- Calculate the Inverse Document Frequency (IDF) for the entire dataset.
- Multiply the TF value of each word by its IDF value to obtain the TF-IDF score for each word.
- Concatenate all TF-IDF scores into a single vector representing the text.

TF-IDF Calculation:

- Load the dataset from a CSV file.
- Iterate over each row in the dataset.
- Preprocess the review text using the text preprocessing steps mentioned above.
- Calculate the TF-IDF scores for each review text using the TF-IDF calculation function.
- Store the TF-IDF scores for each review text in a dictionary, where the key is the review text and the value is its corresponding TF-IDF vector.

3.3 cosine similarity

- Define a function to calculate cosine similarity between two vectors.
- Use numpy arrays to represent the vectors.
- Reshape the vectors if needed to ensure they have the same dimensionality.
- Use the cosine_similarity function from sklearn.metrics.pairwise module to calculate the cosine similarity between the two vectors.

- Return the cosine similarity score, which represents the similarity between the two vectors.

4.Result and Analysis

- Computed cosine similarity between image feature vectors extracted using VGG16 and TF-IDF vectors derived from review texts.
- Higher cosine similarity scores indicate greater similarity between images and review texts.
- Revealed insights into the correlation between visual content and textual descriptions.
- Findings may inform applications such as content-based recommendation systems and image-text matching algorithms.

Using image retrieval

```
Input
Image and Text Query Input :
Image: https://images-na.ssl-images-amazon.com/images/I/81q5+IxFVUL._SY88.jpg
Review: Loving these vintage springs on my vintage strat. They have a good tension and great stability. If you are flo
1/1 [==============================] - 1s 996ms/step
USING IMAGE RETRIEVAL
Image URL: https://images-na.ssl-images-amazon.com/images/I/81q5+IxFVUL._SY88.jpg

Review: Loving these vintage springs on my vintage strat. They have a good tension and great stability. If you are flo

Cosine similarity of images: [[0.9999999]]

Cosine similarity of text: 1.0

Combined Cosine similarity of text: [[0.99999994]]

Image URL: https://images-na.ssl-images-amazon.com/images/I/5145Rvuj71L._SY88.jpg

Review: My oldest son plays and owns many guitars. He asked for a new guitar strap for Christmas. I ordered this one b

The guitar strap has held up very well and my Son says it's comfortable and easy to adjust.

Very happy, that he is happy with his new guitar strap.

I purchased this guitar strap at full price. All opinions and photo's are my own. I am providing this review for anyon

Cosine similarity of images: [[0.9985957]]

Cosine similarity of text: 0.6428146550648325

Combined Cosine similarity of text: [[0.8207052]]
```

```
Image URL: https://images-na.ssl-images-amazon.com/images/I/816L2xBUnKL._SY88.jpg

Review: Thus should be a 5 star review, but sadly, it's only a 3. I bought a cymbal off of eBay that wa

So the finished product is what you see. A far better cleaned up cymbal but still looks like a wreck ar

Cosine similarity of images: [[0.9985768]]

Cosine similarity of text: 0.3389630801889712

Combined Cosine similarity of text: [[0.66876996]]
```

Using Text retrieval

```
USING TEXT RETRIEVAL
Image URL: https://images-na.ssl-images-amazon.com/images/I/81q5+IxFVUL._SY88.jpg
1/1 [==============================] - 1s 802ms/step

Review: Loving these vintage springs on my vintage strat. They have a good tension and great stability. If you are floating

Cosine similarity of images: [[0.9999999]]

Cosine similarity of text: 1.0

Combined Cosine similarity of text: [[0.99999994]]

Image URL: https://images-na.ssl-images-amazon.com/images/I/71nJnXwE9XL._SY88.jpg
1/1 [==============================] - 1s 569ms/step

Review: This bridge is beyond iconic and needed for any Telecaster build. Looks beautiful on any body that you put it on!

Cosine similarity of images: [[1.]]

Cosine similarity of text: 0.9300903833529375

Combined Cosine similarity of text: [[0.9650452]]

Image URL: https://images-na.ssl-images-amazon.com/images/I/71JH6s-YDeL._SY88.jpg
1/1 [==============================] - 1s 570ms/step

Review: I have tried several violin strings and these are the ones I will buy again and again. Over all great even tone. No

Cosine similarity of images: [[1.]]

Cosine similarity of text: 0.92496760541872

Combined Cosine similarity of text: [[0.96248376]]
```

Using combined retrieval

```
USING COBINED RETRIEVAL
Image URL: https://images-na.ssl-images-amazon.com/images/I/81q5+IxFVUL._SY88.jpg

Review: Loving these vintage springs on my vintage strat. They have a good tension and great st

Combined similarity : [[0.99999994]]
Image URL: https://images-na.ssl-images-amazon.com/images/I/71nJnXwE9XL._SY88.jpg

Review: This bridge is beyond iconic and needed for any Telecaster build. Looks beautiful on an

Combined similarity : [[0.9650452]]
Image URL: https://images-na.ssl-images-amazon.com/images/I/71JH6s-YDeL._SY88.jpg

Review: I have tried several violin strings and these are the ones I will buy again and again.

Combined similarity : [[0.96248376]]
```

5. Conclusion

- The study demonstrated the efficacy of leveraging both visual and textual features for content-based similarity analysis.

- Cosine similarity calculations revealed significant correlations between image content and corresponding textual descriptions.

- The combined approach offers promising avenues for enhancing recommendation systems and image retrieval algorithms.

- Future research could explore more advanced techniques for feature extraction and similarity measurement, potentially yielding even more accurate results.