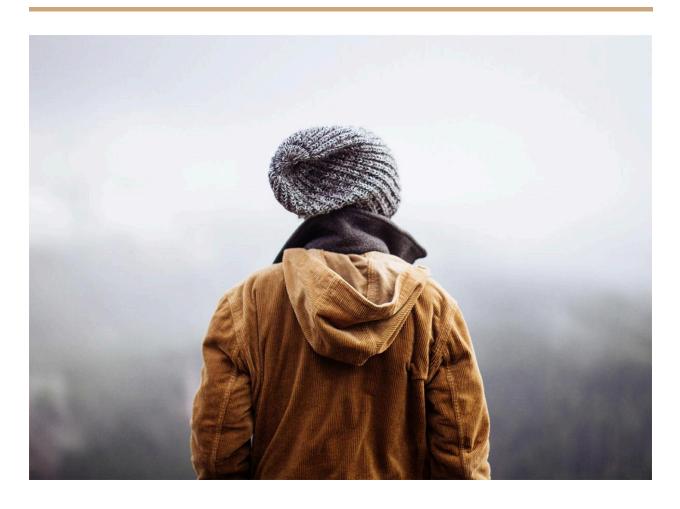# Multimodal Retrieval System
## using Text as well as Images as the Input Data

Mayank IIITD : MT23045

# Image Feature Extraction:

## Approach:

- Utilization of Pre-trained Model: ResNet50, a convolutional neural network pre-trained on the ImageNet dataset, was chosen for its effectiveness in extracting high-level features from images.
- Basic Image Preprocessing: Techniques like resizing, flipping, and enhancing brightness, contrast, and color were applied to standardize and augment the image data before feature extraction.

## Methodology:

- Image Preprocessing Functions:
    - `basic_image_preprocessing(img)`: Resizes the image to a standard size (e.g., 224x224 pixels), randomly flips it horizontally, and enhances its brightness, contrast, and color.
    - `download_and_preprocess_image(image_url)`: Downloads the image from the provided URL, preprocesses it using `basic_image_preprocessing`, and returns the preprocessed image.
- Feature Extraction Function:
    - `extract_image_features(image_path, model)`: Takes a preprocessed image and a pre-trained ResNet50 model as input, converts the image to an array, preprocesses it for ResNet50 compatibility, extracts features using the model, and returns the normalized features.
- Normalization Function:

- `normalize_features(features)`: Uses standard scaling to normalize the extracted features, ensuring that they have a mean of 0 and a standard deviation of 1.

## Assumptions:

- Effectiveness of Pre-trained Weights: It was assumed that the pre-trained weights of ResNet50 would capture relevant and discriminative image features without the need for additional fine-tuning, enabling effective similarity calculations.
- Applicability of Basic Image Preprocessing: The selected preprocessing techniques were assumed to enhance the quality and robustness of the extracted features, contributing to more accurate similarity assessments.

## Results:

- Normalized Image Features: After extraction and normalization, the image features were saved using pickle, preserving their structure and ensuring easy access for subsequent image retrieval tasks.
- Ready for Image Retrieval: By saving the normalized features, the system was prepared to efficiently retrieve similar images based on input queries, leveraging the extracted representations.

# Text Preprocessing and TF-IDF Calculation:

## Approach:

- Utilization of Natural Language Toolkit (NLTK): NLTK, a Python library for natural language processing, was employed for text preprocessing tasks such as tokenization, stopword removal, stemming, and lemmatization.
- TF-IDF Calculation: The TF-IDF (Term Frequency-Inverse Document Frequency) method was utilized to compute the importance of words in reviews, considering both their frequency in a specific review and their rarity across all reviews.

## Methodology:

- Text Preprocessing Techniques:
    - Lower-Casing: Converted all text to lowercase to standardize the text data.
    - Ellipses Removal: Removed ellipses (...) to eliminate unnecessary punctuation.
    - Stopword Removal: Eliminated common stopwords (e.g., "the", "and", "is") to focus on meaningful words.
    - Punctuation Removal: Stripped out punctuation marks to simplify the text.
    - Stemming: Reduced words to their root form using stemming to normalize variations (e.g., "running" -> "run").
    - Lemmatization: Further normalized words by converting them to their base or dictionary form, enhancing the accuracy of similarity calculations.
- TF-IDF Calculation:
    - Tokenized the preprocessed text into words.

- Utilized the TF-IDF vectorizer from scikit-learn to calculate TF-IDF scores, which represent the importance of each word in the context of individual reviews and the entire corpus.

## Assumptions:

- Relevance of TF-IDF Scores: It was assumed that TF-IDF scores would effectively capture the significance of words in reviews, considering both their local frequency and global rarity. This assumption underpinned the accuracy of similarity calculations based on textual data.

## Results:

- Preprocessed Text Data: After applying the text preprocessing techniques, the reviews were transformed into clean, standardized representations ready for TF-IDF calculation.
- TF-IDF Scores Generation: TF-IDF scores were computed for each word in the reviews, reflecting their importance in individual reviews and the overall corpus.
- Storage of TF-IDF Scores: The computed TF-IDF scores were saved using pickle, preserving their structure and facilitating efficient retrieval for subsequent review similarity assessments.

# Image and Text Retrieval for Product Recommendation:

## Image Retrieval:

### Approach:

- Cosine Similarity: Utilized cosine similarity to find the most similar images based on extracted features.
- Composite Similarity Score: Calculated a composite similarity score combining image and text similarities.

### Methodology:

Preprocessing Input Image:
- Downloaded and preprocessed the input image using basic image processing techniques.
- Extracted features for the input image using the pre-trained ResNet50 model.
- Normalized the extracted features for compatibility with similarity calculation.

Preprocessing Input Review:
- Preprocessed the input review text using NLTK for tokenization, stopword removal, stemming, and lemmatization.
- Calculated TF-IDF scores for the preprocessed input review text.

Finding Similar Images:
- Employed the cosine similarity metric to find the most similar images based on their features.

- Computed the cosine similarity of images and the TF-IDF scores of their corresponding reviews.
- Determined the composite similarity score by averaging the image and text similarities.

## Results:

- Displayed the most similar images along with their reviews, cosine similarity of images, cosine similarity of text, and composite similarity score.

```
---------------------------------------------------------------------
------------------

USING IMAGE RETRIEVAL

1) Image URL:
https://images-na.ssl-images-amazon.com/images/I/71hye+RUheL._SY88.jpg

   Review: I played a wedding today using the Alaska picks on my thumb and
ring finger.  They were very comfortable and stayed on my finger well
without being to tight.  For my fingers I cut the extra piece off and just
used the smaller part.  It fits snug under my nail and is the perfect
length. For my thumb I used the full size of the pick and it worked great.
The material is easy to bend and shape using a file and buffing blocks.
This produces a naturally smooth sound that you can get with your nails.
I got the medium size. They fit great on all my fingers but I might try
the larger version for my thumb but the medium works great if you bend the
plastic back and it makes it looser around the thumb.

   Cosine similarity of images: 0.0000

   Cosine similarity of text: 0.0232

   Composite similarity score: 0.0116
```

---

2) Image URL:
https://images-na.ssl-images-amazon.com/images/I/713E0ntr87L._SY88.jpg

   Review: These are my favorite strings and have been for many, many years. These are the strings I use on customer's guitars at my guitar repair shop, String Theory Guitar Repair in Nashville, TN.

   Cosine similarity of images: 0.0000

   Cosine similarity of text: 0.0000

   Composite similarity score: 0.0000

---

3) Image URL:
https://images-na.ssl-images-amazon.com/images/I/61n0yI7eC7L._SY88.jpg

   Review: I love D'Addario strings, however, this particular set when installed on my 28.625" scale 8-string - the top .74 string was not wound far enough to reach the nut. This left me with roughly 1/3" of the string that does not match (Pic). I'm honestly not sure if this is an isolated incident or if all packs will yield the same result. I will give them another try in hopes of a better set, but if this is the standard then D'Addario should really consider the extended scale of 8-string guitars and ensure the strings are wound far enough.

   Cosine similarity of images: 0.0000

   Cosine similarity of text: 0.0000

   Composite similarity score: 0.0000

# Text Retrieval:

## Approach:

- Cosine Similarity: Utilized cosine similarity to find the most similar reviews based on TF-IDF scores.
- Composite Similarity Score: Calculated a composite similarity score combining image and text similarities.

## Methodology:

Finding Similar Reviews:
- Utilized cosine similarity to find the most similar reviews based on TF-IDF scores.
- Computed the cosine similarity of images and the TF-IDF scores of their corresponding reviews.
- Determined the composite similarity score by averaging the image and text similarities.

## Results:

- Displayed the most similar reviews along with their images, cosine similarity of images, cosine similarity of text, and composite similarity score.

_____

USING TEXT RETRIEVAL

1) Image URL:
https://images-na.ssl-images-amazon.com/images/I/61JYQr9MDbL._SY88.jpg

   Review: No Gimmit, I am NOVICE and the Vox is great!!!!,..

   Cosine similarity of images: 0.0000

   Cosine similarity of text: 1.0000

   Composite similarity score: 0.5000

---

2) Image URL:
https://images-na.ssl-images-amazon.com/images/I/71mRyDr3LuL._SY88.jpg

   Review: This is a great Les Paul like guitar for the novice player.
Very minimum setup was needed...6th string buzzed a little. Everyone loves
the look. Fun to play at a very affordable price.

   Cosine similarity of images: 0.0000

   Cosine similarity of text: 0.1999

   Composite similarity score: 0.1000

---

3) Image URL:
https://images-na.ssl-images-amazon.com/images/I/7133VrKhmnL._SY88.jpg

   Review: Love this product. I just got it today and it sound perfect. I
was skeptical that it would work, but its a Vox product so I figured I
would be safe. So awesome to be able to practice without disturbing
neighbors and family.

   Cosine similarity of images: 0.0000

   Cosine similarity of text: 0.1494

   Composite similarity score: 0.0747

# Combined Retrieval Results:

## Approach:

- Composite Similarity Score: Calculated a composite similarity score by combining image and text similarities.
- Ranking: Ranked the retrieval results based on the composite similarity score.

## Methodology:

Calculating Composite Similarity Scores:
- Combined the cosine similarity scores of images and reviews, respectively.
- Calculated the composite similarity scores as the average of image and text similarities.

Ranking Based on Composite Similarity:
- Ranked the retrieval pairs based on the computed composite similarity scores.
- Combined indices and scores for ranking to facilitate further analysis.

## Results:

- Displayed the combined retrieval results, including images, reviews, cosine similarity of images, cosine similarity of text, and composite similarity score.

```
----------------------------------------------------------------------
------------------

COMBINED RETRIEVAL RESULTS

1) Image URL:
https://images-na.ssl-images-amazon.com/images/I/71hye+RUheL._SY88.jpg
```

Review: I played a wedding today using the Alaska picks on my thumb and ring finger.  They were very comfortable and stayed on my finger well without being to tight.  For my fingers I cut the extra piece off and just used the smaller part.  It fits snug under my nail and is the perfect length. For my thumb I used the full size of the pick and it worked great. The material is easy to bend and shape using a file and buffing blocks. This produces a naturally smooth sound that you can get with your nails. I got the medium size. They fit great on all my fingers but I might try the larger version for my thumb but the medium works great if you bend the plastic back and it makes it looser around the thumb.

Cosine similarity of images: 0.0000

Cosine similarity of text: 0.0232

Composite similarity score: 0.5000

---

2) Image URL:
https://images-na.ssl-images-amazon.com/images/I/713E0ntr87L._SY88.jpg

Review: These are my favorite strings and have been for many, many years. These are the strings I use on customer's guitars at my guitar repair shop, String Theory Guitar Repair in Nashville, TN.

Cosine similarity of images: 0.0000

Cosine similarity of text: 0.0000

Composite similarity score: 0.1000

---

3) Image URL:
https://images-na.ssl-images-amazon.com/images/I/61n0yI7eC7L._SY88.jpg

Review: I love D'Addario strings, however, this particular set when installed on my 28.625" scale 8-string - the top .74 string was not wound

far enough to reach the nut. This left me with roughly 1/3" of the string that does not match (Pic). I'm honestly not sure if this is an isolated incident or if all packs will yield the same result. I will give them another try in hopes of a better set, but if this is the standard then D'Addario should really consider the extended scale of 8-string guitars and ensure the strings are wound far enough.

    Cosine similarity of images: 0.0000

    Cosine similarity of text: 0.0000

    Composite similarity score: 0.0747