BURBERRY

LONDON ENGLAND

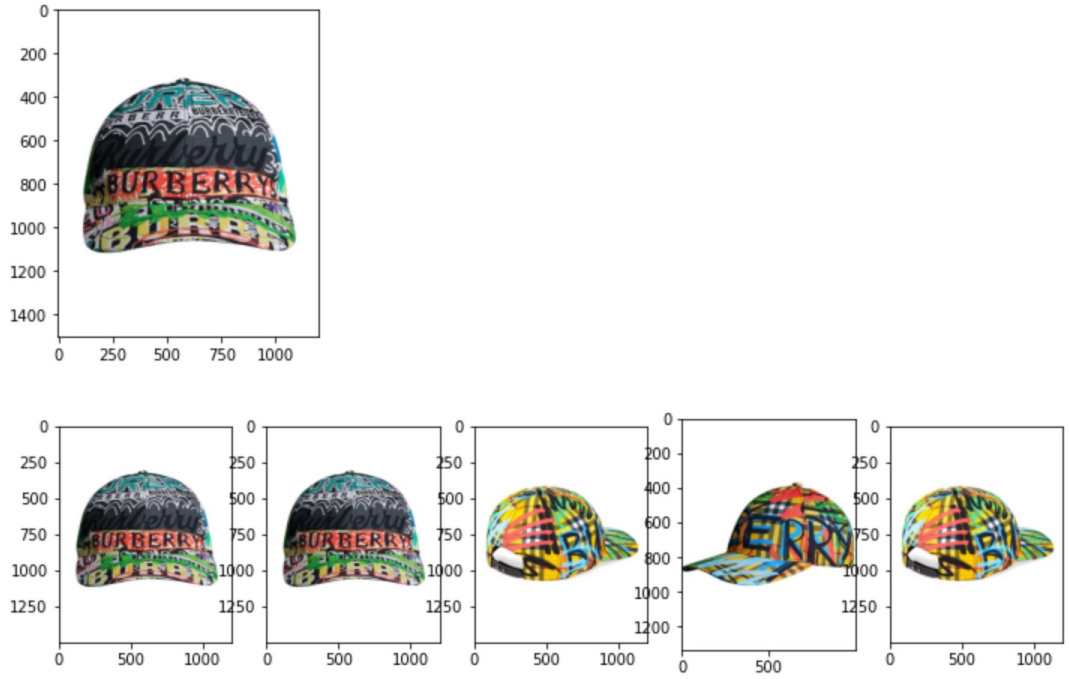
1. PROGRESS SO FAR

A similar image retrieval model based on cosine similarity

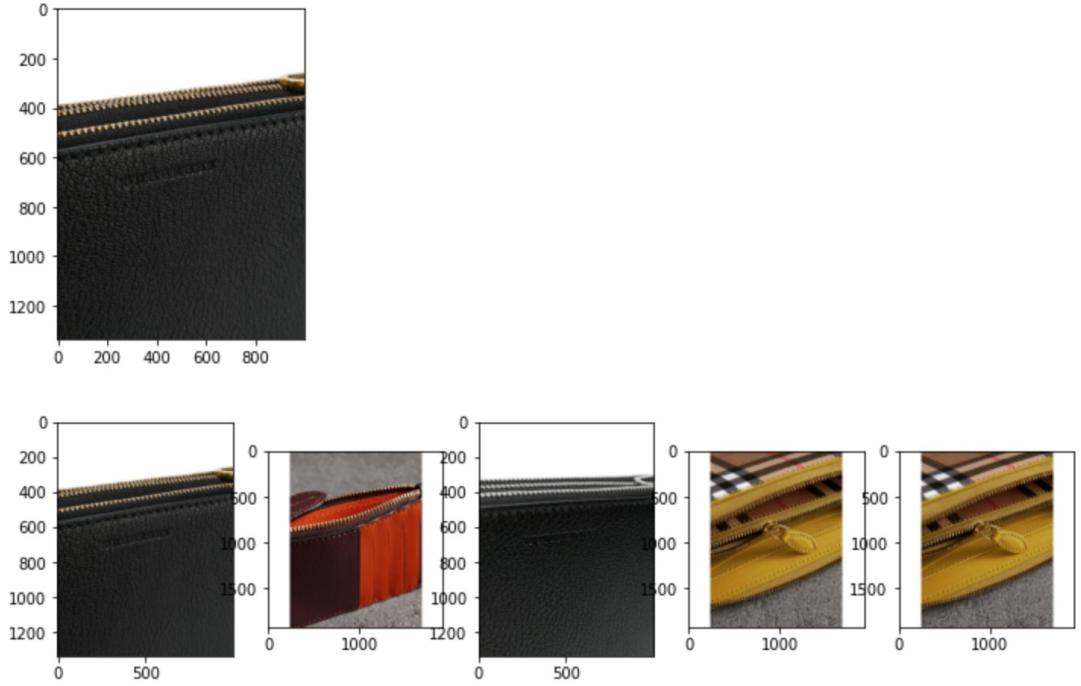
	Text preprocessing	Image preprocessing	Text Embeddings	Image Embeddings
DONE	1. Remove all punctuation/accents/special characters/number/non-alphabetic 2. Lower-casing 3. Lemmatisation 4. Tokenise features	1. Resize and Centre crop 2. Random horizontal flip	1. TDIDF vectorizer	1. Used Resnet18 for generating the embeddings
Not Done	1. Build a dictionary for 2. removing non-meaningful common words 3. Build an escape dictionary for brand name 4. Add part of speech tag	1. Saturation augmentation	1. BERT 2. Add text features from the POS tag	1. VGG16 2. MobileNet

1. PROGRESS SO FAR

An image retrieval model based on cosine similarity (1000 images)



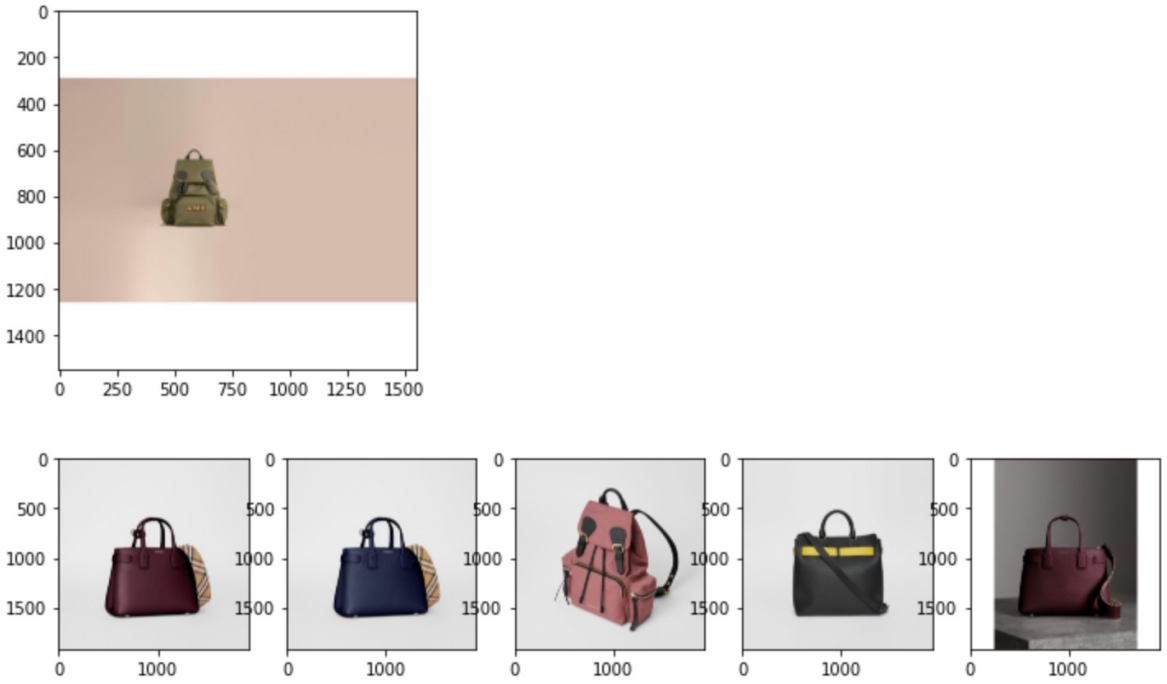
- Seems able to find similar product even with different view



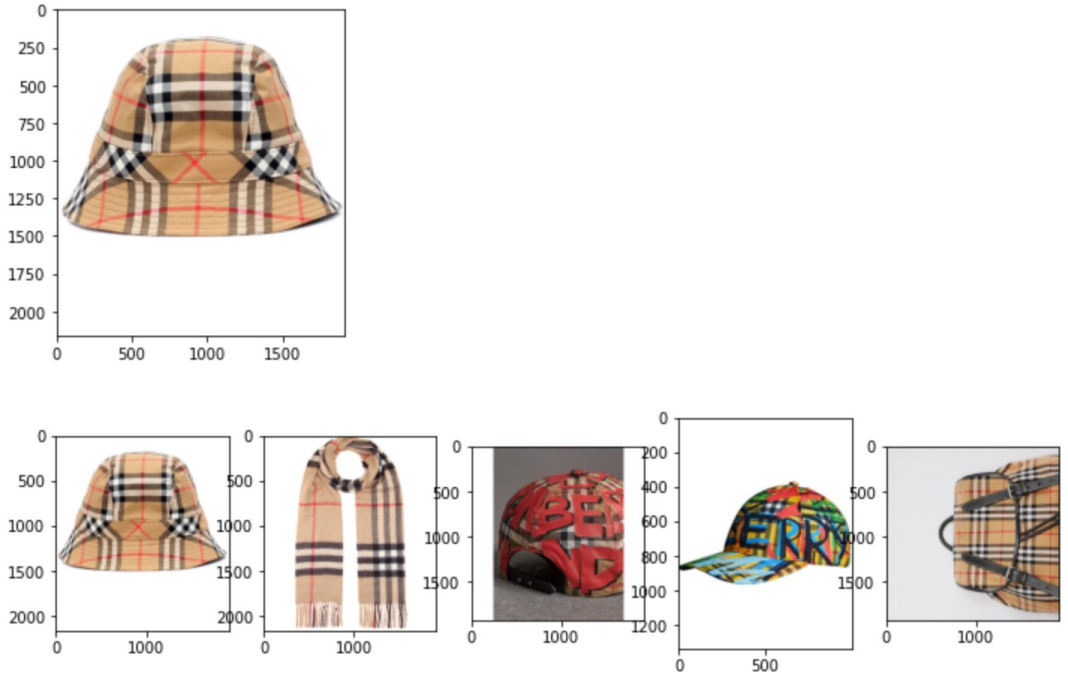
- Fails when the image is a close shoot of product details

1. PROGRESS SO FAR

An image retrieval model based on cosine similarity (1000 images)



- Seems able to find similar product, but same products with different color seems to be hard for this model



- Fails to retrieve products from the same category

2. NEXT STEP

❯ Different Architectures

- Increase datasets size to see if bigger dataset increases the retrieval task performance
- Combine the text embedding with the image embedding
- Use clustering methods or classification to first find the right cluster/category then retrieve similar records
- Siamese network
- Explore different methods of fusion