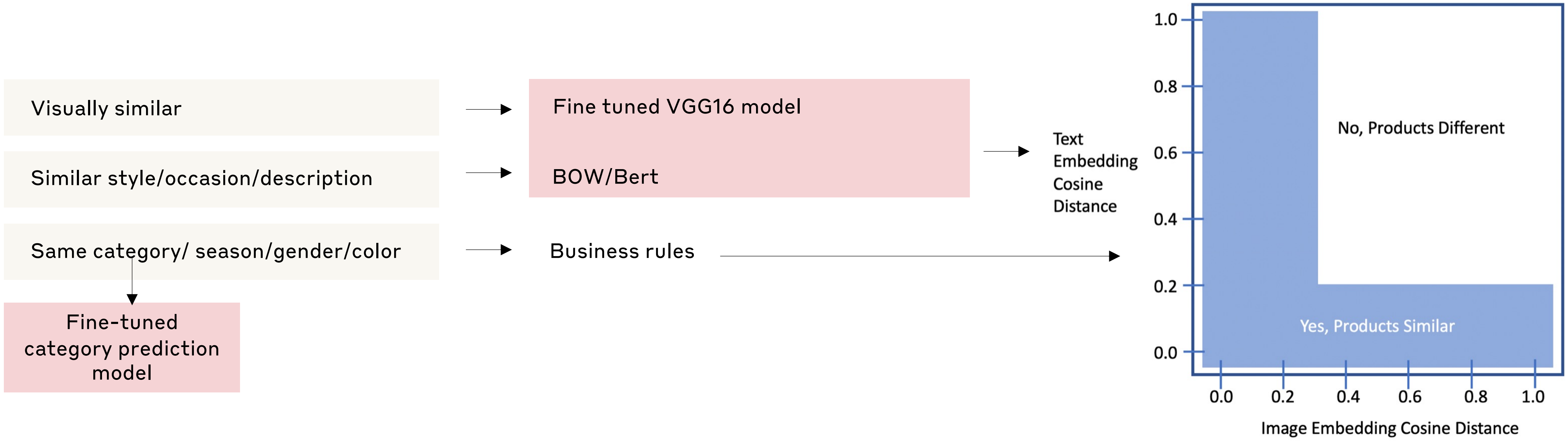
BURBERRY

LONDON ENGLAND

1. IMAGE RETRIEVAL MODEL A

❯ **Combine image and text similarity scores to create a decision boundary**



1. IMAGE RETRIEVAL MODEL B

- Image retrieval model B – Training a CNN on e-commerce data and use as a feature extractor

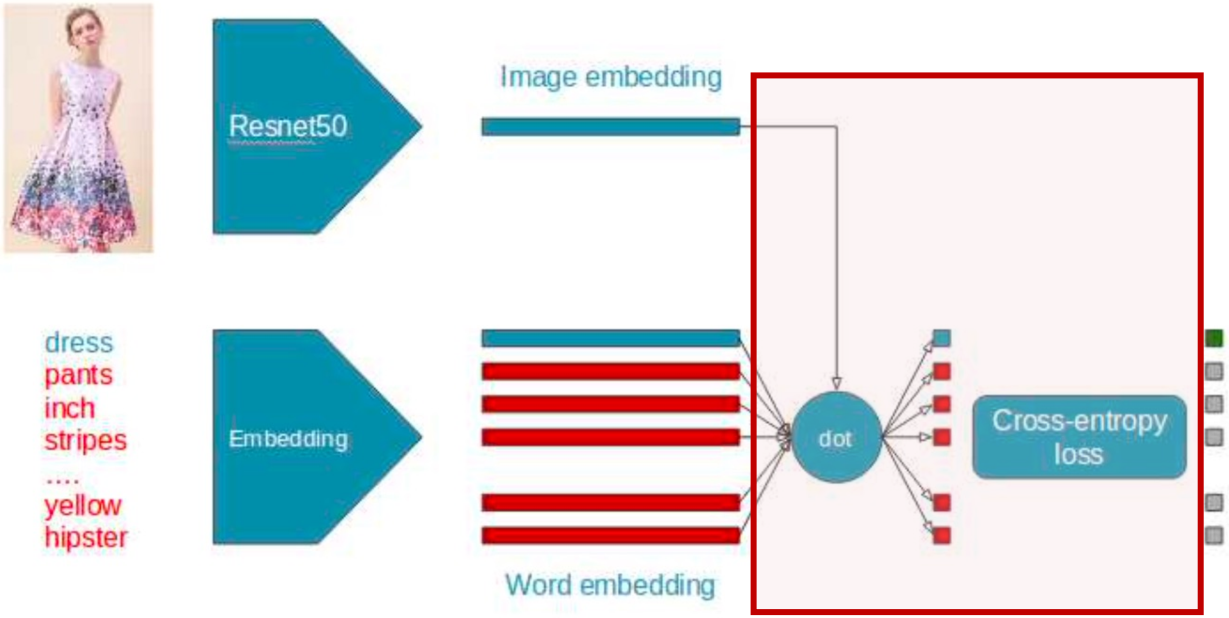


Figure 2. Training of our model: predict one label, picked from the bag-of-words description, from an image. Both image and words are embedded before being coupled in a dot product. We output a probability for each word in the vocabulary.

2. PROGRESS SO FAR

Updated and built a new prediction model

- **Built bag category prediction model using Burberry in-house bag image and level 5 label data**

  Used the model to predict labels on the competitor product images using Burberry product hierarchy. We want to force the final model to retrieve products in the same label. E.g., Query image is a tote bag, the retrieved results will only include products that has the same predicted category. This means we need to have a decent prediction accuracy otherwise the retrieved results is less accurate.

- **Updated VGG16 image embedding extraction model using Burberry still life bag images**

  Better retrieval results on within the edited competitor dataset, images contain same products are retrieved together with similar scores unlike previous results. But not significantly better for Burberry-competitor unseen image retrieval, most likely due to the limited competitor image database only 4096 images including duplicate products for now.

- **Update the retrieval functions with Faiss**

3.PERFORMANCE COMPARISON

Prediction base performance on the EDITED data

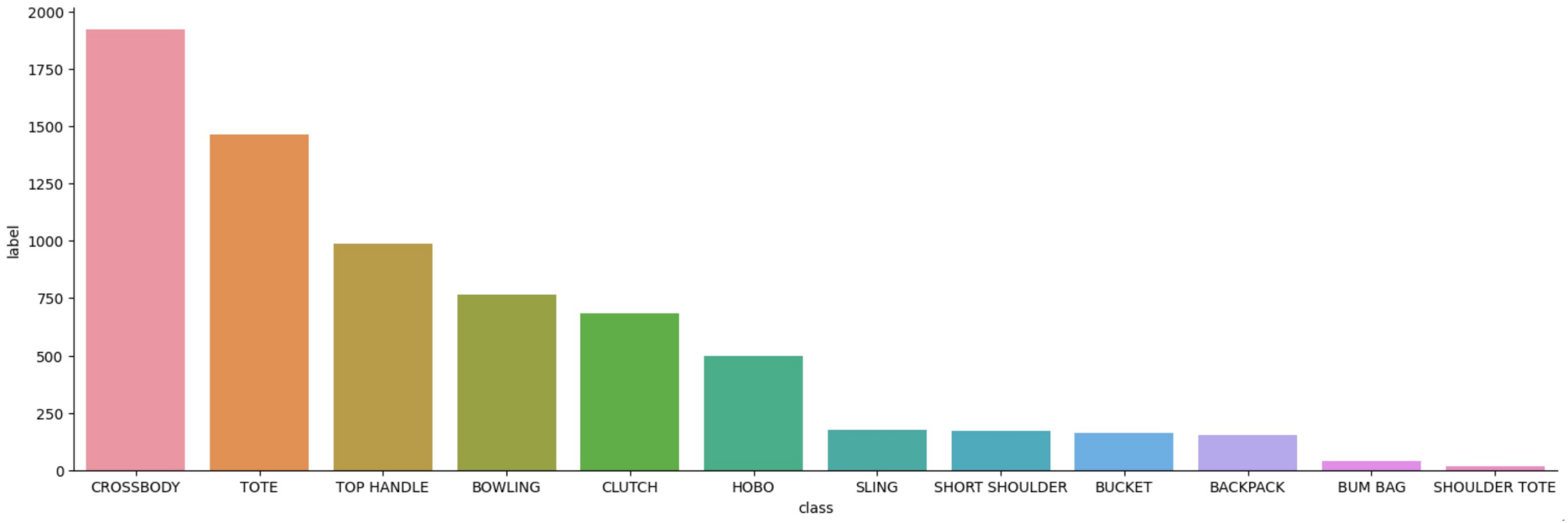
-Category prediction base performance on the raw data

Data used to train model	Model	Test accuracy (Burberry data)	Top1 accuracy (Edited data)	Top3 accuracy (Edited data)	Top5 accuracy (Edited data)
• Uniformly sampled still life data with shot type <5 (4200) • Only clean product images	V1 (overfitted)	0.922	/	/	/
• Sampled 1 image still life data for all unique material id (3800) • Only product images with clean background	V2 (very imbalanced)	0.885	1.62%	/	/
• Sampled non still life data for all unique material id (7900) • Contains both images human model and products images, unable to separate between the two	V3 (very imbalanced)	0.709	6.91%	19.6%	23.9%

4. CLASS DISTRIBUTION

Class distribution of the data used to train model v3

-Category prediction base performance on the raw data



5. NEXT STEP

Different Architectures

- Working on retrieval function using Faiss
- Resample the data to correct imbalance
- Use Faster R-CNN to crop out ROI to build bigger dataset