Deep Neural Network Ensembles for Extreme Classification

Yichen Pan, Yuxuan Sun

The main task in this project is about the so called Extreme Classification, where we need to deal with multiclass involving an extremely large number of labels. Specifically, the training dataset contains 12,118,359 images (6,924,452 products), validation set contains 252934 images (147,758 products), and testing set 3,095,080 images (1,768,182 products). In total, there are 5,270 categories, and the final goal of the project is to correctly predict products into these 5,270 categories.

First, we experimented with several deep neural network models, including ResNet50, Se-ResNet50, ResNet101, InceptionV3, XceptionV3, etc. with a limited training epochs, to test single model performance and set the baseline for further experiments. Then, we tried Test Time Augmentation (TTA) and network ensembles, together with prediction aggregations to increase performance. Finally, we tried pseudo labeling to further fine-tune the trained neural network in a semi-supervised fashion, hoping to further generalize the model to the testing set and improve the model performance on product-wise prediction (i.e. not simply image-wise prediction).