

Pairwise Compatibility Prediction

Data Preparation

- The txt and json files for train, test and val should be present under polyvore outfits directory.
- The python script reads the txt file and corresponding JSON file. For each ie .. for test, train, val the scripts creates csv files and then inputs the dataloader. The test.json file has no impact and will not be used in the test loader. The test loader is a split from the train data.
- So no manual intervention or processing is required.
- Please ensure test.json is present under polyvore_outfits directory.
- The intermediate CSV files that have been generated have also been attached as train_compatibility.txt, test_compatibility.txt, valid_compatibility.txt

Approaches

- There are different ways to approach a binary classification problem. I chose Siamese network and Modified pre trained resnet50 by adding FC layers with Sigmoid output which is most commonly used for binary classification. I have included the code for both the approaches. However, limited by time and resources, I was able to partially implement Pre-Trained resnet50.
- The training of the siamese network requires many Convolutional layers/blocks. But the training time for 23 Convolutional layers increased drastically and I couldn't run it on P3 instance even with a lowest batch size.

Implementation

- For pairwise classification with siamese network, we feed forward two images separately and calculate the Euclidean distance based on margin defined. This is used to predict the compatibility.
- With resnet50, the modified pre trained network took a lot of time to load and was not feasible to train in the given time.

Performance Tuning

- I have incorporated Data Augmentation and Learning Rate Step Scheduler. I could see slight increase in accuracy but step scheduler was not required for convergence

Model plot

- This couldn't be generated because torchsummary didn't accept multiple inputs for forward pass. Hence I couldn't generate model summary.

Python script files

- The data.py and utils.py have been renamed as data_compatibility.py and utils_compatibility.py. Please rename to run the python files.