Creating train and test sets from pre-processing script. Load the training and testing datasets in finetune.py to get the results.