

Prateek Narang

SIY177546

MS(R), SIT

IIT Delhi

ELL-881 Deep Learning

12st October 2018

Paper Title: "Learning General Purpose Distributed Sentence Representations via Large Scale Multi-task Learning"

OVERVIEW

In this project, completed the following tasks for training a Model.

- 1) Natural Language Interference (NLI)
- 2) Constituency Parsing
- 3) Natural Machine Translation (NMT)

PURPOSE

Purpose was to learn generalised sentence representations (embeddings) by performing multiple different tasks as mentioned in the paper Model uses a shared RNN Encoder for all the above tasks.

TIME

Training was performed on GPU Only(with dropouts)

- NMT Training - 1 epoch ~ 2 hrs 20 mins
- Constituency Parsing Training - 1 epoch ~ 2 hrs 25 mins
- Machine Translation - ~ 12 hrs+ (and counting)

EVALUATION RESULTS

1. The model was evaluated on the tasks 10.1, 10.2, 10.3 and 10.4 as mentioned in the paper.
2. 5 nearest neighbours of each word were generated as given in 'nearest_neighbours.txt' file.
3. The NMT Training was taking a lot of time, due to approaching deadline, only a subset of dataset around 70,000 batches each of size 8 were used for Machine Translation task. The NLI and Const. Parsing completed 1 epoch over entire dataset.

Evaluation Results - 1 (After Complete Training(1 epoch) after NLI Task with dropout)

{'CR': {'ndev': 3775, 'devacc': 68.73, 'ntest': 3775, 'acc': 66.97},
'SST2': {'ndev': 872, 'devacc': 65.02, 'ntest': 1821, 'acc': 59.97},
'MPQA': {'ndev': 10606, 'devacc': 73.14, 'ntest': 10606, 'acc': 73.08},
'SICKEntailment': {'ndev': 500, 'devacc': 64.0, 'ntest': 4927, 'acc': 66.19},
'SUBJ': {'ndev': 10000, 'devacc': 71.42, 'ntest': 10000, 'acc': 71.67},
'SICKRelatedness': {'yhat': array([3.0914448 , 4.66949518, 2.1914208 , ..., 3.14491807, 4.06996782, 4.68601826]), 'ndev': 500, 'devpearson': 0.6721336727622759, 'ntest': 4927, 'spearman': 0.6615137412170515, 'pearson': 0.6858121088199186, 'mse': 0.5523532818876562},
'MRPC': {'f1': 80.09, 'ndev': 4076, 'devacc': 69.55, 'ntest': 1725, 'acc': 68.75}, 'SST5': {'ndev': 1101, 'devacc': 30.52, 'ntest': 2210, 'acc': 31.9},
'MR': {'ndev': 10662, 'devacc': 60.12, 'ntest': 10662, 'acc': 59.55}}

Evaluation Results - 2 (After Complete Training on 3 Tasks with dropout)

{'MPQA': {'devacc': 74.03, 'ntest': 10606, 'acc': 73.35, 'ndev': 10606},
'MR': {'devacc': 59.41, 'ntest': 10662, 'acc': 57.97, 'ndev': 10662},

'SST5': {'devacc': 29.16, 'ntest': 2210, 'acc': 31.09, 'ndev': 1101},
'SICKEntailment': {'devacc': 65.2, 'ntest': 4927, 'acc': 63.59, 'ndev': 500},
'SUBJ': {'devacc': 77.11, 'ntest': 10000, 'acc': 77.4, 'ndev': 10000},
'MRPC': {'ntest': 1725, 'devacc': 70.27, 'f1': 80.76, 'acc': 68.81, 'ndev': 4076},
'CR': {'devacc': 68.98, 'ntest': 3775, 'acc': 67.02, 'ndev': 3775},
'SST2': {'devacc': 61.93, 'ntest': 1821, 'acc': 62.77, 'ndev': 872},
'SICKRelatedness': {'pearson': 0.6682779742396917, 'spearman': 0.6491135642552158,
'devpearson': 0.6460745590455536, 'mse': 0.566765278613459, 'ndev': 500, 'ntest': 4927,
'yhat': array([2.43179922, 4.37165187, 2.92893252, ..., 2.44479083, 4.54899041,
4.67519554])}}

REFERENCES & CITATIONS

1. The pre-processed(parsed) dataset for constituency parsing, and evaluation scripts were provided by Anshul Mittal (2017siy7545), Modified the evaluation script to suit the requirements as per my model. Anshul also helped in many insightful discussions over some concepts and common errors.
2. ELL-881 course repository was also very useful for learning concepts like RNN, Dropouts, Setting up Data Pipeline. Some of the code snippets have been modified and used in the project.

<https://github.com/vineetm/ell-881-2018-deep-learning/blob/master/project1/project1.md>