

NAME: Lee Zheng Yao Daniel

EID: dzl237

CS378 NLP Project Check in

Progress thus far:

- Reviewed relevant given literature to get an understanding of the project's relevant concepts and dataset
- Dataset Selection: I have decided to go with option 2, and choose a question and answering dataset, SQuAD (Rajpurkar et al., 2016) to use. It is a reading comprehension dataset, where the answer to a question can be found in the associated passage
- Model Selection: ELECTRA-small as recommended, which has the same architecture as BERT
- Set-up: Successfully set up provided started code, and trained the started code on the SQuAD dataset4
- Baseline model to compare to: Original BERT model. It has trained on the SQuAD dataset before with SoTA results, serving as a good baseline for us. Link to paper: <https://arxiv.org/abs/1810.04805>
- Preliminary data exploration: Looked through some randomly selected examples of the SQuAD dataset, to get an understanding of what kind of questions etc. are used.
- Preliminary formulation of a possible data artefact hypothesis: I suspect that there may be some bias in the way questions are phrased, which makes the model able to guess an answer more easily without proper reading comprehension. For example, questions starting with "who", "where", etc. will hint that there is a high likelihood of the answer being a person's name or a location respectively.
- Decision of possible fix plan: Focused learning--curriculum learning. Reviewed paper at https://ronan.collobert.com/pub/2009_curriculum_icml.pdf on curriculum learning, by Joshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. This paper introduces the concept of curriculum learning, where we train our model on a sequence that gradually increases in difficulty. The authors argue that this can help models converge faster and find better local minimas compared to traditional methods. They demonstrated the effectiveness of this, with one example being language modelling

Plan going forward:

- Applying Dataset Cartography technique to identify the different subsets of data to understand what is easy to learn, hard to learn, and ambiguous. I can then analyze these with my prelim hypothesis in mind to get a better hint of the possibility of it
- Fixing plan: Use focused learning in line with our hypothesis to try and mitigate dataset biases. (Curriculum learning) I currently plan to design and set up a curriculum learning strategy plan on the SQuAD dataset after applying Dataset Cartography, and then compare the performance of this against the original baseline BERT model
- Fix evaluation: Evaluate effectiveness of the fix given

- Consideration of further improvements: Depending on the results of my fix implementation and evaluation, I will consider if I want to apply additional techniques to improve the performance
- Final Report and analysis