Building a QA system (IID SQuAD track): BiDAF with Answer Pointer

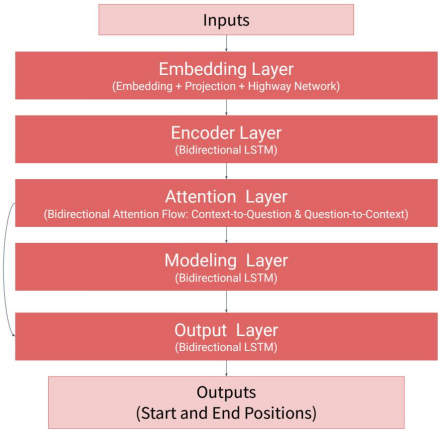
Catherine Wang

Computer Science Department, Stanford University

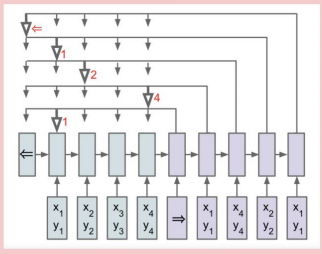
Stanford Computer Science

Problem: The question answering task involves returning the correct answer given a question and a paragraph of context. This is a central task in NLP that is used to evaluate machine comprehension of natural language. The SQuAD dataset has a particular characteristic where the answer is a section of text within the given paragraph.

Background: The baseline for this project is the Bidirectional Attention Flow (BiDAF) model, minus character-level embeddings. The model takes as input two arrays of word indices for the question and the context paragraph and runs them through the following architecture:



Methods:
The Answer Pointer layer, based on the Pointer Network architecture, replaces the output layer of the baseline model and conditions the end position on the start position.



The original Pointer Network architecture

Character-level embeddings for the inputs were also implemented in the embedding layer.

Analysis:
- Both character embeddings and Answer Pointer led to slight decreases in performance
- More hyperparameter tuning may help
- Answer Pointer alone may not be enough for significant improvement (no ablation study done in original paper)
- Answer Pointer may be biased towards producing answers over N/A

Conclusions:
- Importance of elements aside from model architecture
- General best practices when engaging with research
  - Reading papers that are similar to or contributed to the current paper
  - The usefulness of ablation studies

Experiments: Baseline | Character Embeddings | Character Embeddings + Answer Pointer