QANet Analysis: Default Project (non-PCE)

Timothy Le

Problem: Given a passage and a query, predict the start and end indices in the passage to answer the question, or predict “No Answer”

Data: The SQuAD is used to approach this problem

- **QANet Architecture**
  - Implemented Encoder, Model, and Output layers from QANet model:

- **QANet diagram from QANet paper**
- **Attention equations from “Attention is all you need” paper**

- **Example output from baseline w/ character embeddings**

- **Effects of making the QANet model more complex**
- **F1 and EM on last training Epoch**

  - Dev NLL on last training epoch
    - Large drop in loss from adding residual connections
    - Loss is too high, compared with baseline (3.05)

  - Baseline: BiDAF model w/o character embeddings
    - Train: Dev NLL: 3.05, F1: 61.53, EM: 58.24
    - Test: NLL: 3.06, F1: 61.27, EM: 58.46

  - Best score: BiDAF baseline w/ character embeddings
    - Test: NLL: 2.94, F1: 62.29, EM: 59.10

- **Conclusion and Future Directions**
  - Residual connections throughout the encoding layer are vital in reducing the loss (Dev NLL)
  - Increased complexity in one component of the model should have corresponding increased complexity in related components of the model
  - Future: Tune parameter such as dropout rate to lower training loss and investigate layers since the fullQA model only predicts “No Answer”

- **References**