Building end-to-end QA model with reward based objective

TEAM REINFORCE (Prashant and Eti)

Error Analysis on SearchQA

Question	Predicted	Gold Answer	Reason for error	%error
party u singer also plays young lady named hannah		Miley Cyrus	Wrong answer span First para talks only about miley (no mention of hannah)	
signature appetizer p f chang 's chicken cups vegetable	Crisp lettuce	lettuce	-	
ground coffee brand 's best part wakin'	folgers	folger's	Preprocessing error (Answer not present in document)	

Interesting Observations

1. Interestingly, in majority cases first occurence of the answer does not seem to correspond to the right context of the question

2. To answer few questions, we need external knowledge ie. question and passages are not related
   Eg : question : 4 x 12 ?

   answer : 48
   Answer is present in the passage but context is different.(passage talks about construction)

3. In few cases, we also observed that questions don't make any sense without metadata
   Eg : question : barbara undershaft
   answer : major barbara
   Category for this question is "Play heroines". Without this information, it is not possible to predict the answer.

Tasks in pipeline

1. Write the code for passage ranking(baseline Rl model) (✓)

2. Complete and successfully run the code for distant supervision

3. Integrate BiDAF model with RL model (✓)

4. Hyperparameter tuning

Timeline

March 03 - March 12	March 13 - March 30	March 31 - April 15	April 16 - April 30
~~Implementation and training of own BiDAF model in dynet~~	~~- Perform experiments on SQuAD dataset~~	~~- Analyze model on SearchQA Dataset~~ ~~- Implement the passage ranking functionality~~	-Run the complete model(BiDAF + passage reranking -Error Analysis and further improvement