Development Log

July 17, 2018

Deployed Model to Slack

July 15th Notes

Sunday Notes

Steps taken

1. Limited to just baseball
2. Tokenized Sentences
3. Removed Hyperlinks from text
4. Increased scoring threshold
5. Changed the vocabulary size
6. Changed the tokenization scheme (BPE to regex)
7. Limited text to < 100 characters

Possible Courses of Action

1. Continue to Tweak the chat engine
   1. Add bigrams phrases to excluded phrases
   2. Mandate English words
   3. Stopwords
   4. Remove Numbers
   5. Remove ALL Special Characters
   6. Try sentDex tweaks
   7. Add data
2. Use another LSTM chat engine
   1. I have installed https://github.com/tensorlayer/seq2seq-chatbot