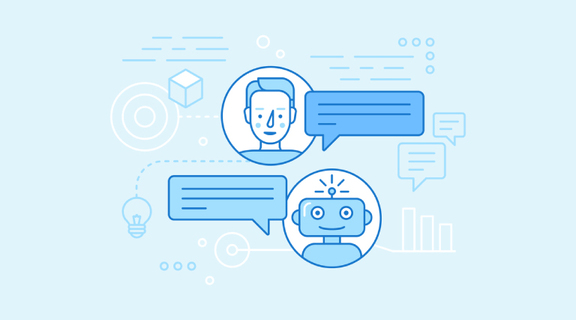
|  |  |  |
| --- | --- | --- |
| https://www.alfavita.gr/sites/default/files/styles/default/public/peloponnisoos.jpg?itok=hT65wkfU | **MSc IN DATA SCIENCE**  **UNIVERSITY OF PELOPONNESE –**  **NCSR “DEMOKRITOS”** | https://www.alfavita.gr/sites/default/files/styles/default/public/images/2018/11/ekefe.jpg?itok=Ll4iYuTX |

Deep Learning

Assignment 2: Build a Chatbot!



A report by:

Georgios Touros (dsc18023)

January 2020

## **Experimental Setup and Data Pre-processing**

We downloaded the original, unprocessed corpus[[1]](#footnote-1) and followed the instructions that were attached in the zip file to pre-process it. Our goal was not to use any of the metadata that are provided, but instead, to simply construct pairs of questions and answers that could be used to train our chatbot.

|  |  |  |
| --- | --- | --- |
| Parameters | Model | Conversations |
| Min Line Length = 2  Max Line Length = 10  Min Word Frequency = 1  Conversations Sample = 30,000  Questions Used: 13,131  Vocabulary Size: 6232  LSTM Hidden Layers = 200  Batch Size = 100  Epochs = 45  Wall time: 30min 38s |  |  |
| Min Line Length = 1  Max Line Length = 12  Min Word Frequency = 2  Conversations Sample = 30,000  Questions Used: 27,750  Vocabulary Size: 7036  LSTM Hidden Layers = 200  Batch Size = 100  Epochs = 45  Wall time: 1h 25min 7s |  |  |
| Min Line Length = 1  Max Line Length = 12  Min Word Frequency = 2  Conversations Sample = 30,000  Questions Used: 27,799  Vocabulary Size: 6,994  LSTM Hidden Layers = 400  Batch Size = 100  Epochs = 45  Wall time: 2h 22min 58s |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

1. https://s3.amazonaws.com/pytorch-tutorial-assets/cornell\_movie\_dialogs\_corpus.zip [↑](#footnote-ref-1)