Deep Learning Project

Group member: Yinjia Liu & Chenwei Cao

November 25, 2020

* Introduction

A chatbot is a software that provides a real conversational experience to the user, here for our project, we build an open domain chatbot which can generate a response as the name implies.

* Tools and Datasets

The tools we used are Keras on Python, and google colab; the dataset we used are collected from Kaggle.

* Data Overview

the datasets we use are human responses and robot responses each contains 2363 entries records.

* Data Pre-processing

We used regular expression to clean our data, then we grouped human response with robot response as following code shows:

Text

Description automatically generated

* Model Selection

We selected seq2seq model as our training model.

Diagram

Description automatically generated

Figure 1

As figure 1 shows above, The seq2seq model also called encoder-decoder model which use long-short term memory for text generation from the training sets. It predicts a word given in the user input and then each of the next words is predicted using the probability of likelihood of that word to occur.

* Model Evaluation
* Conclusion
* References