Chatbot design using Policy Gradients

Mini-Project Description:

This objective of the project is to design two chatbots(Agent A and Agent B) using policy gradient based training strategies so that they can handle natural conversations between them. The project will use LSTM based encoder–decoder architecture trained on a conversational database .(See Reference Code and Dataset in next slide) The RL based fine-tuning will be done on top of the trained LSTM encoder-decoder . The details of the RL based fine-tuning is available in the **reference paper** and the **reference code**. We will refer to the RL based fine-tuned model given in the reference code as the 'baseline' version. The project is expected to achieve two high level objectives on top of the baseline version

High Level Objective 1:

The baseline version has used certain reward formulation and achieved some success in having prolonged conversations between Agent A and B as compared to the initial LSTM encoder-decoder model trained with only the conventional supervised way. Design and implement additional strategies that can further prolong meaningful conversations between Agent A and Agent B.

High Level Objective 2:

Design training techniques that will give two separate personas for Agent A and Agent B. Agent A should be showing more positive behavioural characteristics (e.g. empathy, optimism) , while Agent B should be showing more negative behavioural characteristics (e.g. arrogance, cynicism)

Chatbot design using Policy Gradients

Reference Paper: Deep Reinforcement Learning for Dialogue Generation

Reference Code 1: https://github.com/pochih/RL-Chatbot

Reference Dataset 1 :https://convokit.cornell.edu/documentation/movie.html

Reference Code 2 https://github.com/Ls-Dai/Deep-Reinforcement-Learning-for-Dialogue-Generation-in-PyTorch.

Reference Dataset 2 : https://paperswithcode.com/dataset/dailydialog

*. You are free to choose between Reference Code 1/ Reference Dataset 1 and Reference Code 2/ Reference Dataset 2. Alternatively any other implementation of the reference paper also can be used as your reference code.

Instructions

- The project package should contain code, trained model and the technical report. Each team needs to upload only one package in the LMS.
- Maximum four members are allowed in a team. .
- The technical report should contain details of project members (name, Id) in the beginning as well as a project group Id.
- The technical report should contain details of reward formulation and training strategies used to achieve the high level objectives. Subjective conversational comparisons with the baseline version and your method should be provided to showcase the results.
- A script should be provided to generate conversation between the Agent A and Agent B. The script should take in the first conversation (for Agent A) from the user. Details on how to run the script should be given in the technical report.