



Watson

A crime solving sidekick

Team Members:

Shyam Kumar Nair(skn42)

Talin Sai Upputuri(tu50)



Introduction

- The conversations on Reddit are the closest we have to unfiltered human dialogues.
- In recent times, there have been many cases that Reddit has solved.
- One such case:-
 - Eight years ago, A person was repeatedly found posting notes on the back of his chair and sometimes on his doors, which is neither his handwriting nor was there any proof of break-ins. This incident was posted to Reddit.
 - One person in the comments had just put out his thoughts and intuition about the problem, which actually solved the case.
 - It was found that the person had been suffering from Carbon Monoxide poisoning because CO was leaking into his apartment, which was later detected by the CO detector.
 - Years later, this thread was turned into a Podcast.



Introduction

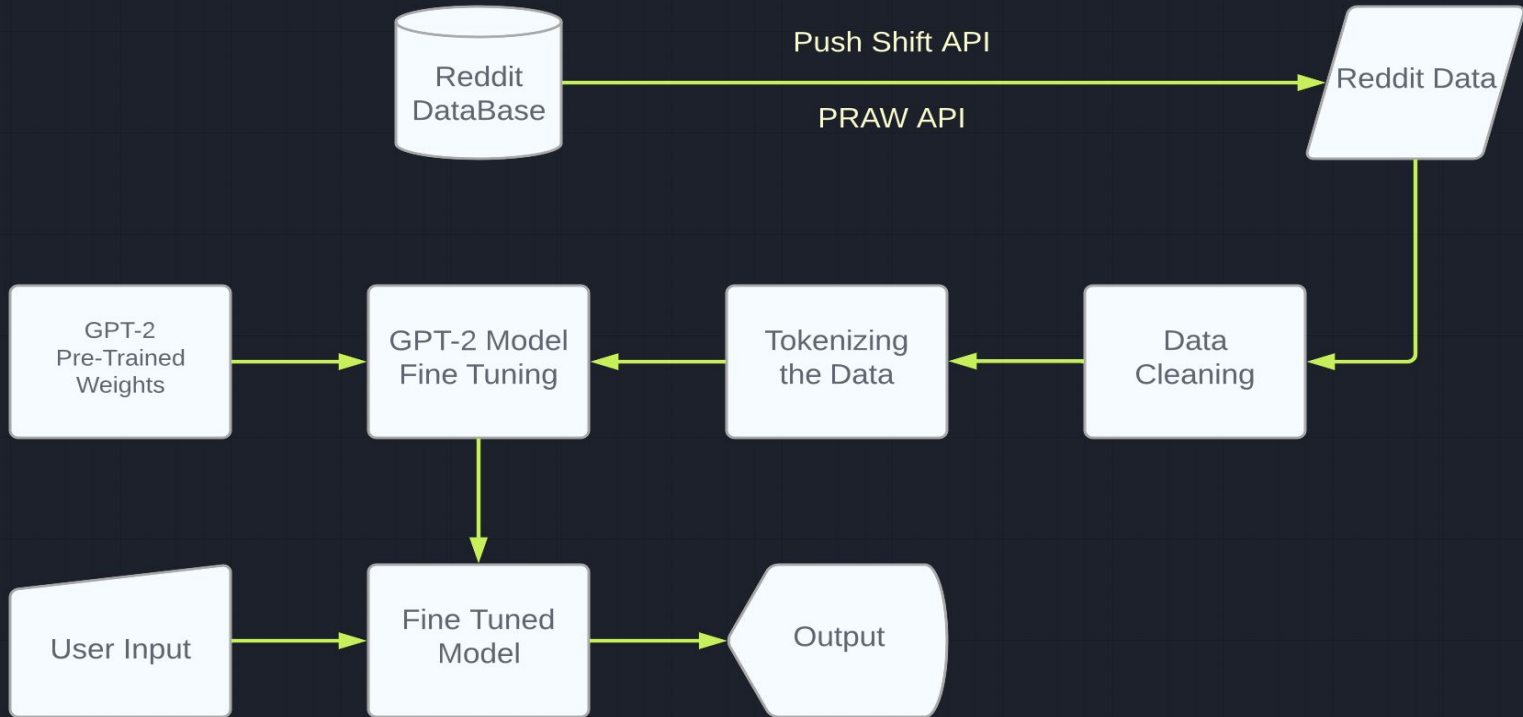
- Developing a Crime Chatbot using Reddit data
- There have been cases that could not have been solved without a small amount of help from Reddit.
- By providing a convenient and accessible way for individuals to report crimes and get help, A crime chatbot can potentially make a real difference in people's lives.
- We Planned to solve the Problem using deep learning to train the model with the Reddit data and then answer the questions related to crime.
- Our baseline Model used GPT-2, which was trained on fewer data and produced garbage values.



Data Collection

- Two different methods were used for the Data extraction:
 - PRAW(Python Reddit API Wrapper)
 - PushShift API
- Both have limitations but combined, a large amount of data was collected.
- Initially, 20-30k rows of data were collected for the baseline model.
- Later using both APIs, this was increased to ~400k
- Cleaned the data by removing the punctuations, stop words, and threads deleted by the users.
- The data was collected starting from when the subreddits were created
- The dataset was divided with an 80:20 split for training and testing.

Architecture



Structure of Input Data

input_ids

<bos>

<Question>

Question

<Reply>

Reply

<eos>

token_type_ids

<Question>

<Reply>

labels

<msk>

<msk>

<msk>

<msk>

Reply

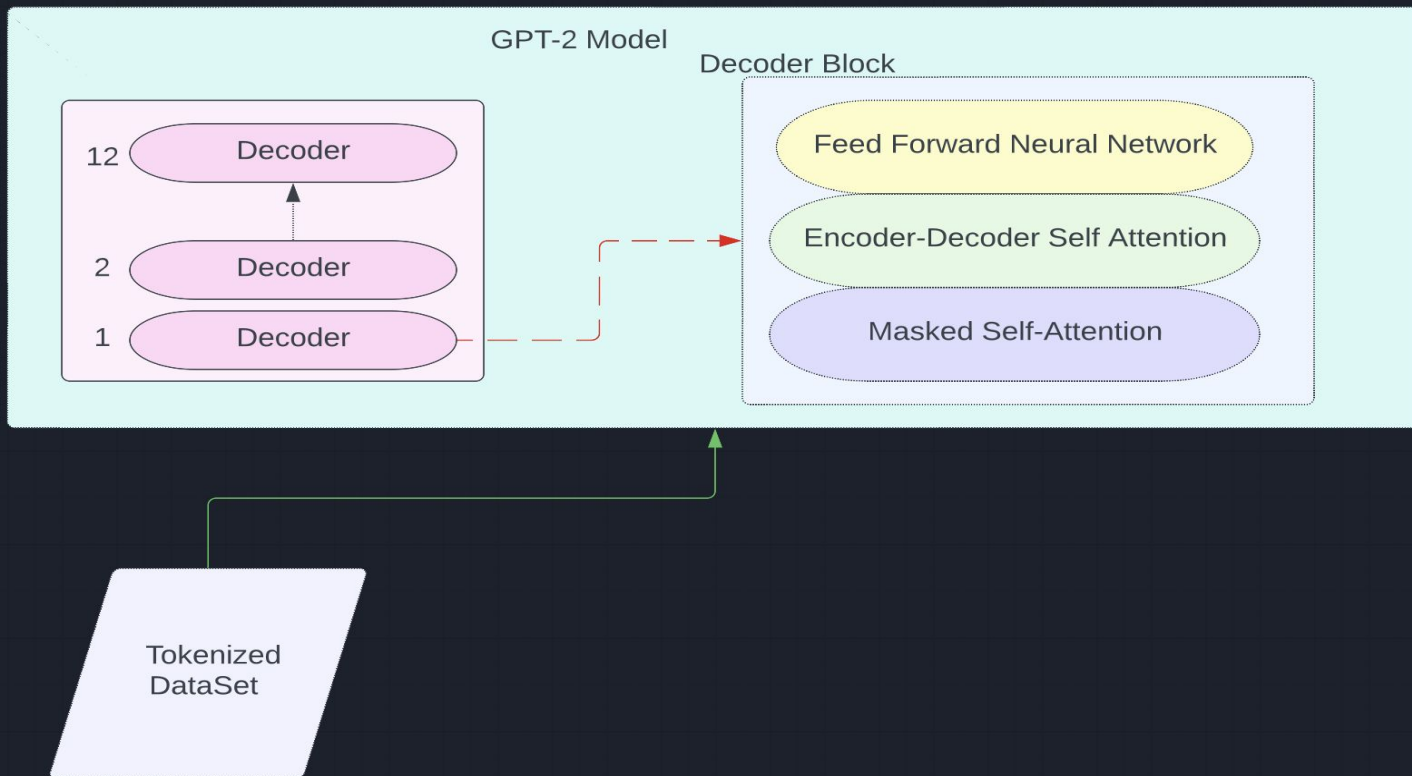
<eos>

Intro to GPT-2



- Currently, the most widely used chatbot is the ChatGPT which uses the following version of GPT-2, i.e., GPT-3.
- GPT-2 is a type-of language processing modeling developed by OpenAI.
- GPT-2 Model is a pre-trained model that uses a neural network architecture called a transformer.
- The pre-trained GPT-2 model was taken from HuggingFace.
- The link to the Pre-trained model:- <https://huggingface.co/gpt2>

GPT-2 Architecture





Trainer API & TrainingArguments

- The Trainer class provides an API for feature-complete training in PyTorch for most standard use cases.
 - TrainingArguments
 - Datasets
 - Model
- The TrainingArguments is a parameter of the Trainer API which helps us provide other arguments, such as
 - Learning Rate
 - Optimizer
 - Devices to use



Results

- The output of the architecture is a set of probabilities for the next possible word.
- The loss function for this Training is Language Modelling Loss.
- The following is our loss
 - Training Loss - 2.96
 - Validation Loss - 2.83



Demo

User >> Who is Tiffany

Watson >> Valiante?She is missing. She came home to a stranger.

User>>> What is Tiffany Valiante Theory?

Watson >> I agree that the Tiffany Valiante case is extremely strange because it's the most likely scenario. The evidence points to foul play, and this person was clearly in extreme need of treatment.

But what about the other theories in the article? If she had just been stabbed she'd have been dead much sooner (a lot later) and not survived long after the stabbing.

Just the facts and contextThere's a reason we do not believe that the stabbing happened from the second episode. At the end of the episode, Tiffany seems to have had her phone stolen and was



Limitations & Future Changes

- High Variance
- Structure of Data and Computational Capacity
- Elimination of unnecessary posts (use of scores)
- More advanced architectures available
- Adding publicly available case files to the database

Thank You

