**Shakespeare GPT**

**Abstract**

One paragraph summarizing the entire paper

**Introduction**

One of the difficulties of implementing cutting-edge technology is determining its best use case. The goal of the authors research is to provide a practical implementation of a Generative Pre-Trained Transformer (GPT) that meets the desired use case of the end user. The most cutting-edge example as of the time of writing is ChatGPT3. Which uses unsupervised learning to train on vast amounts of data and can then be asked questions. The result of the authors research will be a GPT, trained on all of Shakespeare's writing that will allow for similar use as ChatGPT3, however, it will be trained on a unique set of data and will employ Reinforcement Learning from Human Feedback (RLHF) method. The result was the user is able to ask imperative text questions such as "Describe the character Henry the Fifth", with a coherent answer such as "Henry the Fifth was King of England and fought in the Battle of Agincourt." The query returned was timely, accurate, and truthful response that answers the intent of the user’s query. Accurate answers are the paramount priority of the tool. Answers to queries will be returned with a relative probability (between zero (0) and one (1) with one being 100% accurate) that the returned answer is correct.

**Machine Learning and Generative pre-trained Transformers**

ML is a method for teaching machines how to process data without explicitly defining every possible response. A variety of different algorithms can be applied to a dataset to allow the computer to return the results in the users desired format. There are two primary ways that this occurs, supervised and unsupervised learning. Supervised learning takes labeled data to train models which can then predict unlabeled data based on the training the model received. Unsupervised learning is the training of models with an example of the rules which allows the model to look at data that is unlabeled and manipulate it based on the set of rules it was provided.

**Supervised Learning**

Supervised learning is the simpler form of ML and is the most common. Supervised learning is used when the input and output are both known quantities. For example, if a

**Unsupervised Learning**

**Generative pre-trained Transformers**

Generative pre-trained Transformers (GPTs) are language models that are used to produce human-like text that will respond to prompts from users. GPT’s responses are determined by the dataset it was trained on. For example, the most popular version of a GPT, OpenAI’s GPT-3, was trained on 175 billion parameters of open-source data including textbooks and Wikipedia[[1]](#endnote-1). Because of the profusion of information that GPT-3 was trained on it does not require any further training (Radford, A. Narasimhan, K. Salimans, T. Sutskever, I; 2020). However, for data that is not available in datasets that GPT-3 had access to, it cannot accurately provide answers. Datasets that GPT-3 does not have access to could include proprietary business information or even classified government datasets.

-description of previous iterations of GPT

Previous iterations of GPT

**Methodology**

The tool will utilize mixed methods to maximize the existing qualitative and quantitative algorithms to produce the most efficacy for our use case. The qualitative method will use the Reinforced Learning from Human Feedback (RLHF) algorithm to provide human input to train the reward function which will assist in determining how the tool will respond to textual inputs. The resulting reward function will then be applied in an unsupervised quantitative way to train on the corpus of data, in this case, the entirety of Shakespeare's written works. (Add description of Quantitative method and transition)

**Qualitative Method**

One of the requirements for GPTs to function is the training of the language model. The training of the language model can be conducted in a variety of ways with varying degrees of efficiency or tuned for certain use cases. In the case of our tool, it was trained with the Reinforcement Learning from Human Feedback (RLHF) method. In this instance, all of Shakespeare’s works and some critical essays were tokenized by word, character, and subworld to allow for fine-tuning and then a cross-entropy calculation to was applied which allowed Shakespeare's GPT to return English or similar words, in this case, it should reproduce language like that of the Bard. (Insert discussion of training parameters and size etc.)

The generation of English or similar text allowed us to train a scalar reward model that rewards the model the most for the most correct answer and continuously decreases amounts until the least correct answer is achieved. We then reached a decision point in which we had to determine if they should continue training the model based on human inputs or allow the computer to train the model based on the inputs that it had already received. We chose to allow the computer to train the reward model based on the human inputs for n number of n's. We then conducted another set of human interactions with the model and then conducted another iteration of computer training. The result was applied to our evaluation data to test if we had overtrained the model. The resulting reward model was then applied to the totality of the works of Shakespeare and allowed to train.

**Quantitative Method**

Conclusion:

-why did we use them together?

-why is ours different than everyone else’s | what did we do differently that makes our widget better than OpenAI’s widget

**Results**

We need to compare the results to something

-how have the results from our widget proved that we have improved (incrementally or generationally) on the works of other scholars

**Discussion**

-Discuss our findings

-Conclude with future avenues of research

**References**

Radford, A. Narasimhan, K. Salimans, T. Sutskever Improving Language Understanding by Generative Pre-Training

1. OpenAI, git hub <https://github.com/openai/gpt-3> [↑](#endnote-ref-1)