An Overview and History of GPT-2 and OpenAI

This technology review will go over the history of OpenAI and their incredible progress with their unsupervised natural language processing model GPT-2. OpenAI began as an open-source non-profit venture, founded by billionaires Elon Musk, Sam Altman and a couple others who wish to remain anonymous. The goal of the organization was to promote widespread academic development of a robust and friendly AI that would promote the welfare of society. The program was eventually part-acquired by Microsoft and later transitioned to a for-profit model. Some deeper motives for the creation of OpenAI were to prevent what Elon Musk described as “the biggest existential threat to humanity”, this threat being Artificial Intelligence.

The bigger goal of OpenAI was to be a proactive measure to this possible threat but critics claim that OpenAI could become exactly what it sought to destroy. While OpenAI has countless ventures in AI based activities such as robotic metalearning, video game bot creation, and AI debating, this review will focus on its primary product which are its generative models, namely GPT and GPT-2. GPT stands for Generative Pre-Trained Transformer. Before we get into how it works, let's dive into what it can do. One of the coolest functions of GPT-2 is its ability to generate text. On command GPT-2 can generate entire passages to prompts and and even answer questions. Of course it is not perfect, it was found that many prompts not frequently found on the internet can cause the model to output ridiculous text which is not understandable, but for the most part this function tends to be the most popular use of GPT. The accuracy and state-of-the-art technology GPT uses has scared some who believe that it can be misused to generate fake news, or generate short text-segments (like tweets for

example) that are indistinguishable from a human-generated output. Other less-used functions of GPT-2 are summarizing text, and translating text between languages. Now to get into the technical side of GPT-2. The model is surprisingly actually unsupervised. This means it trains without any notion of a correct input/output pair. In otherwords it is able to learn at a much quicker rate and in a way that avoids the problems of overfitting. Actually, training a language model at this scale in a supervised environment would be virtually impossible because at that scale there is no dataset that perfectly mirrors correct/incorrect outputs. In term's of scale GPT-2 trains 1.5 billion parameters. This is quite a lot but the successor to GPT-2 trains on a scale of 100 billion parameters. The transformer based nature of the model allowed for more parrellelization, so the model was able to train on a custom made dataset called WebText which was generated by crawling through reddit links.

In conclusion, GPT-2 uses perhaps some of the most state-of-the-art deep learning methods we have seen thus far in the field. Because of its unique features it was able to utilize the growth in computational power through its efficiency in learning.

Sources:

https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf