Daniel Cufiño, Tochukwu Okwuonu
COMP 646

Project Proposal: Prompt-based ASCII Art Generation

I. Basic Idea

The general idea of our project is to generate ASCII art based on a text description of the desired ASCII art. For example, if you entered the prompt "Give me ASCII Art of Pikachu surfing" it would do exactly that. We have two main approaches for accomplishing this task but we aren't sure which might be the most suitable. However, both of these ideas will require the same training inputs.

II. What is an ASCII image?

ASCII art is an art form that creates images using ASCII characters from the computer.

Examples:

III. What makes our AI unique?

Based on our research there is no AI that can accomplish this task, most of the ones take in text and convert it into different fonts using ASCII characters. None of them create the images. If you also try using existing language or vision models like DALL-E-2 of GPT-3 the results of generating an ASCII image based on a prompt are usually quite poor.

IV. Dataset:

We're still searching for viable datasets to use for this problem. One issue is that we may want our ASCII art-generating problem to be domain-specific to improve success. As a result, we're still trying to identify datasets that meet this criterion. One such dataset is this dataset found on HuggingFace which contains less than 1000 (Pokemon image, text description) pairs. Another possible dataset would be to create our own dataset of (emoji, text description) pairs. This might be a time-consuming process however, it may be possible to automate it though through the use of some sort of Image2Text Model but we aren't immediately aware of one such type of model to use for that purpose.

V. Resources:

If we decide to use a GPT-3 model provided by OpenAI, we will likely use their API to train and evaluate our fine-tuning. The pricing for use of these model can be found here

Inputs:

1. **To train the model:** (image, image text description) pairs. We will explain how exactly this data is used.
2. **To generate results from the model:** a text description of the desired ASCII art.

Outputs:

A string of text characters that represent that ASCII art corresponding to the prompt provided.

Approach 1:

The first approach we thought of is to use a Text2Image Model like DALL-E-2 to generate an image corresponding to the provided prompt. With this generated image we can then "ASCIIfy" the image. It is possible to generate ASCII art for an image in multiple ways. Firstly, there are pre-existing DEEP Learning models on the internet that convert an image to ASCII, however we are not sure we can count on its robustness for general images. There also exists methods for "ASCIIfying" any image using Python which can be used to partially generate the predicted image. The general idea of this approach is to train a model to minimize the difference between the predicted ASCII art and the "Ground Truth" ASCII art (which is just the "ASCIIfied" ground truth image generated using Python). Another approach may be to use a GAN with the "ASCIIfied" predicted image as a starting point in the hopes that the GAN will learn to imitate the

Approach 2:

The second approach we thought of is directly training a model similar to GPT3, with text and ASCII images. It would make it solely a language model, removing the need to convert the descriptive text into images. We're still ironing out the details for this approach and some advice would be appreciated. For example, we're not exactly sure what loss metric we'd use or which language model exactly we'd like to use. Our first instinct is to fine-tune the GPT3 model

but an issue with this is that we're not sure how the pricing works for fine-tuning the model and it may also require a large corpus of training data to be effective.

Evaluating Success:

We plan on evaluating success by putting it to use. We will have other students at Rice use it to create ASCII images of their choice and have a satisfaction rate from 1 - 10. Using this score we will then fine-tune our model to get the highest scores possible. Our goal will be a satisfaction of over 9. Depending on the success of our project, we're also interested in possibly turning into a Widget / App on the app store that can be used on iMessage to be able to easily send custom ASCII art to people through SMS.