Redefining Online Communication: Unique Artificially Created Emojis
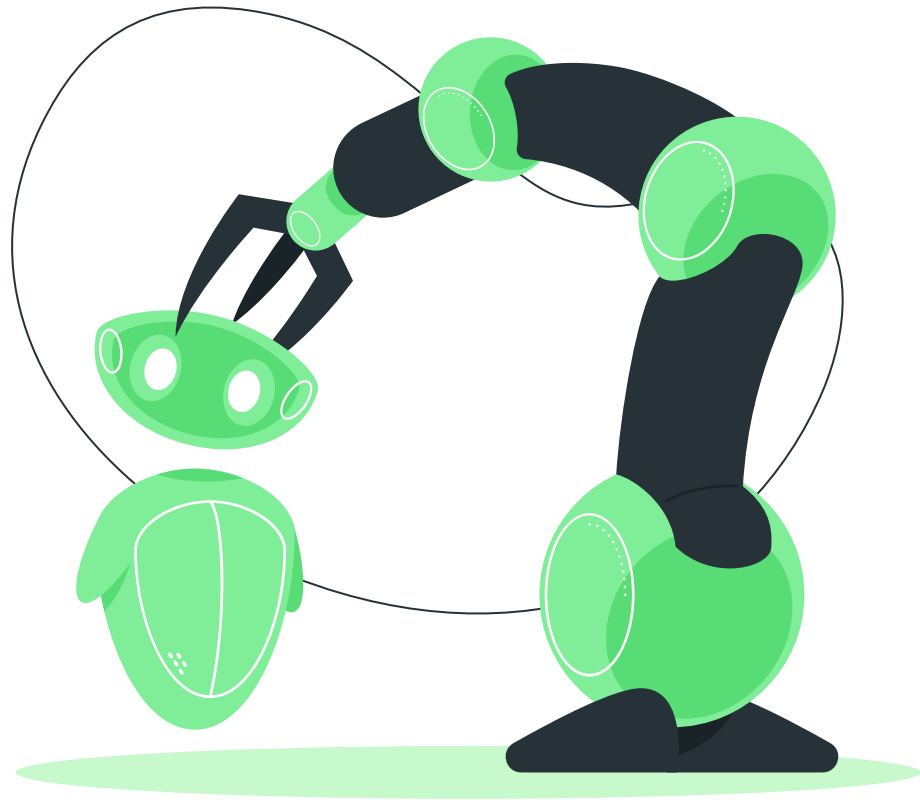
COMP 646 Course Project

Daniel Cufiño
Tochukwu Okwuonu

Motivation

1

Assessing Image-Generating Capability of GPT Models

We aim to assess whether LLMs are capable of producing text meant to represent an image. Currently deployed LLMs either do not seem capable of such a task or have not been trained to do so.

2

Bolstering Creative Communication

We hope to create a tool that can be used by individuals to allow them to express themselves online more freely.

Problem / Data Setup

Data Source

- {Image, text caption} dataset of captioned emojis from numerous companies (Apple, Twitter, Microsoft, etc...)
- Used a pixel-by-pixel, brightness-based algorithm to convert images to text
- Formatted caption and ASCII art as prompts.
- Data is now ready for training and modelling.

Modeling

We use **GPT-2 Medium**, a large transformer decoder model because of its ability to generate new text. This is because it was trained specifically to generate the next word in a sequence.

- GPT-2: 350 million parameters
- Released by OpenAI
- 1024 Maximum Input Token Limitation

We fine-tune this GPT-2 model in our project with the data collected and cleaned from the previous slide. Since the images were converted text, they can be used as input to the model.

Experiments

- We first train a single sample from our data on the model to see if it can learn the task.
- Takes roughly 1000 iterations to memorize the a single training sample.

- Due to time and computational ability contrainstraints, we realized it was unfeasible to fully train our dataset of 14,000 samples.
- Even using the best available GPU on Google Colab Pro (A100 GPU w/ 40 GB GPU RAM), we estimate training to have taken ~855 hours or 35 days to complete.
- We instead train on a limited subset of samples from our dataset to prove the model can generalize and generate unique emoji ASCII art.


Figure 4. Result of training our model on a single sample from our dataset of numerous iterations

Results & Future Work

We received satisfactory results given the GPU RAM limitations we had:

- Based on the campus-wide survey, we conducted our model was given a satisfaction rating of **8.375 out of 10** with a high of 10 and a minimum rating of 5. The results of the survey is given in the chart on the top right.
- Visually, as displayed on the bottom right, the results of our experiments and model is consistent with what we hypothesized in the beginning of the process.
- In the future, we think that the current created model can be used to train larger datasets on a machine with larger computing power.
  - We also think the model can emulate platform specific emojis if prompted to.

Performance Rating by Count
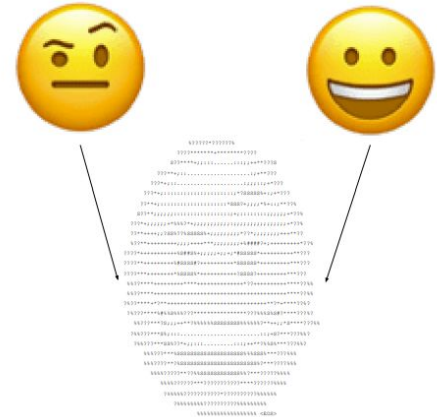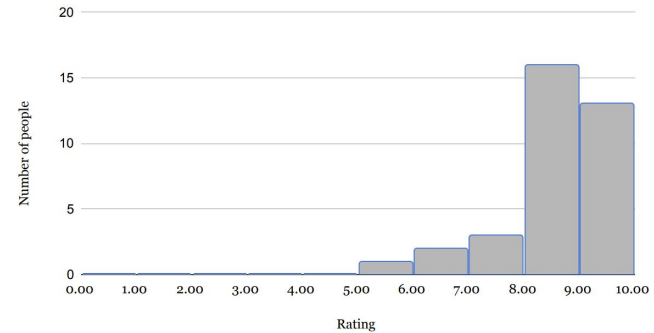
Thanks!

We've both learned so much this semester and are very thankful for your engaging lectures and assignments. We are also very helpful for all the feedback and support you've given us along the way!

Alternative Resources

Find more illustrations like these on **Stories by Freepik**