Mini-Project 1

Noah Foster

CSCI 2952N

I really liked this miniproject and found it is relevant in many ways to my own research and even my final project from deep learning! I will walk through the project and discuss some of the ways that I think my implimentation stands out and some of my justifications and logics for the code I wrote. Hopefully this ordering is relatively similar to the Jupyter Notebook I submitted. I should note that I tried to keep the envirnonment as close to the original as possible, but I did have to make some changes to get the code to run on my machine. I also wanted to use the library Spacy as well as a couple others. Anything not included in the email this report surely arrived in is probably in this GitHub Repository. In the name of making all the code work, I have also renamed gpt3-sandbx to gpt3sandbox.

Task 1: Zero Shot

Note: The project guidelines mention CIFAR-10, while the provided notebook uses CIFAR-100. I use both just for fun.

The logic for this part is very similar to that of the provided Interacting with CLIP notebook. I am using a Macbook Air for this project and thus struggle to run anything except ResNet-50, so further experiments are solely done with ResNet-50. I tried the other but mainly just found the difference in time. The core logic is that CLIP is trained to embed images and captions so that the cosine similarity between the two is high when they have similar content. So we build a caption that corresponds to each class and see which classes align best. I tried a couple methods of building captions and found some interesting things. The format of caption that works best for CIFAR-10 is not the one that works best for CIFAR-100. It is also interesting that in both cases, ending the sentence with a period does not hurt and in fact helps despite making the tokenization a little more strange.