Isaiah Martinez
Joycelyn Tuazon
Prof. Lord
Comp 542
September 18, 2023

Project Proposal

Our project is on image classification of cats. We will be feeding an Unsupervised Learning Machine Learning Algorithm images of cats and non-cats in order to see how effective the computer is at determining the likelihood of an image containing a cat given an unseen image that may or may not contain a cat.

The dataset we have decided upon is:

- Image Classification of CAT vs NOT A CAT (DOG/PANDA): https://www.kaggle.com/datasets/ashishsaxena2209/animal-image-datasetdog-cat-and-panda

The unsupervised algorithms we will most likely use are amongst the following:
- KMEANS - known # clusters (3 based on the different animals)
- ISODATA - unknown # clusters (based on individual features of animals being looked at)

Additional Preprocessing steps that need to be done to better model creation:
- Resizing of images to be standard/uniform
- Gaussian blurring to highlight subject within the photo

We intend for the evaluation of the model to increase when the preprocessing is done. Perhaps we can compare the models for the additional preprocessing for the images vs the given images in order to see if the additional preprocessing gave a sufficient boost to model performance.

Our goal is to find the Maximum likelihood of the image containing a cat. We are focusing on Classification of the subject of the image (whether the subject is a cat or not). We hope for this project to better our understanding of Machine Learning as well as image classification techniques used in industry.