

MAIS 202 Data Selection Proposal

Project Idea: Create a webapp that can take a selfie as input and replace the face in the photo with an emoji corresponding to that individual's expression.

Dataset: https://github.com/muxspace/facial_expressions

- I chose this dataset because there are approximately 13,000 photos, each of which is mapped to an expression.

Methodology:

1. Data preprocessing:
 - a. The dataset I chose is highly feasible as it's from Kaggle. However, I will have to discuss with execs on how to train nearly 50MB of data considering that during the hackathon my team and I ran out of RAM constantly when working with only a fraction of this amount of data.
2. Machine learning model:
 - a. I want to classify expressions in photos and present this classification as an emoji
 - b. I would likely use a convolutional neural network (CNN) using the Keras library. I'm not sure if there are other alternatives that can do the job better, so I'll keep an eye out for possible options in upcoming lectures.
3. Evaluation Metrics:
 - a. Confusion matrix and accuracy/precision-recall/logistic loss
4. Final conceptualization
 - a. I will create a simple webapp.