Preliminary Results

by Sarah Zhou

**1. Problem Statement**

The objective of this project is to create a machine learning model that will train on a mix dataset of deepfake and non deepfake videos to predict if a video has deepfake content or not.

**2. Data Preprocessing**

I chose the set from [Deepfake Detection Challenge](https://www.kaggle.com/c/deepfake-detection-challenge/data). It is a large dataset of 470GB. I am not doing data augmentation. There are 400 videos. The one label per video is either REAL or FAKE, used for binary classification and feedback.

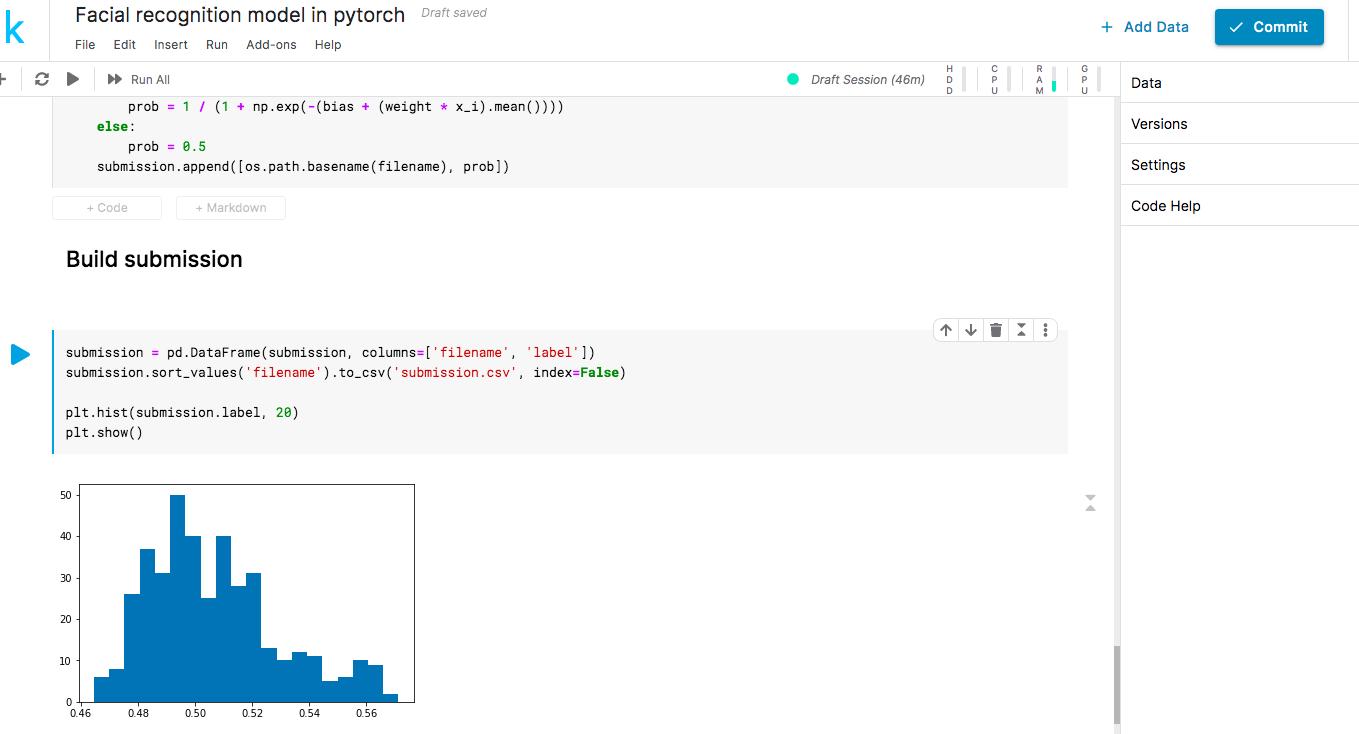
**3. Machine learning model**

The pretrained facial detection model chosen here is MTCNN and Inception Resnet models. We then use a package named Facenet by Pytorch. It is still a CNN, thus I am still on the same plan as stated in my proposal. Neural networks are the only possible way to process videos and perform deep learning, thus the I will finetune and try alternate versions of CNNs. The dataset was already split by the competition provider, thus I did not make any decisions to alter it.

It was challenging to find possible models to train this humongous dataset. As predicted, the dataset requires a lot of space and power to process, thus I decided to directly work in Kaggle notebooks to facilitate the process. In brief, after a long research turning in circles, two TPMs suggested to directly try out implementations from the Kaggle community.

**4. Preliminary Results**

From the results :



The model seems to be wildly underfitting for a binary classification problem. The clusters are too continuous and there is not a good separation of the output. The small rage of 0.46 to 0.56 cannot be used to distinctly

The training time was around 45mins.

Although a face detection pipeline was successfully created, it was not effective.

**5. Next Steps**

I will either tune this model or try another one. I am considering to use pretrained models released by **FaceForensics++**. I researched and a baseline model for deepfake using **dlib package**. It seems like a significantly better jump off point then the previous model[[1]](#footnote-1). The preliminary results from this model gave a clear binary separation.

1. <https://www.kaggle.com/robikscube/faceforensics-baseline-dlib-no-internet/output> [↑](#footnote-ref-1)