Problem Statement

This project whether training neural networks using metrics better suited for unbalanced datasets such as ROC-AUC, balanced accuracy, and F1 combined with post-training decision threshold tuning can lead to higher test accuracy than neural networks trained solely using accuracy. I additionally aim to examine how class distributions affect this result

Dataset

My project will examine the fashion MNIST dataset (https://github.com/zalandoresearch/fashion-mnist). This is a dataset consisting of black and white images of articles of clothing belonging to 10 different categories. The multiclass nature of this dataset allows the construction of multiple binary classification problems as well as multiclass classification problems. I will additionally be able to resample the data from this dataset to construct various class distributions to examine the effect that class distributions have on results. This allows me to evaluate how class distributions affect results. evaluating how findings extend to the original multiclass setting.

Methods, Techniques, and Technologies Under Consideration

I will use TensorFlow to implement various CNN and CNN-LSTM classification models. Training will use a standard loss such as cross-entropy, and will vary its use of techniques like class weighting. I will use metrics such as accuracy, ROC-AUC, balanced accuracy, and F1 score for model training, but I will only use accuracy when adjusting the decision threshold post-training. Additionally, decision threshold adjustments will be made using a validation set to prevent data leakage. Depending on the results of each binary classification problem, I will consider dropping samples to further vary the class distributions. However, if dropping samples results in too small of a dataset for meaningful learning, I will additionally consider using an additional dataset for examining results under varying class distributions.

Products to be Delivered

The final deliverable will include a python notebook, a final report, and a video. The python notebook will document all of my work and summarize my results. The final report will further discuss my methods, results, lessons learned, and will suggest future directions of this work. Finally, I will present my findings in a video to be presented in class.