Machine Learning II
Final Project #1 Proposal
Nov. 21, 2017
Brent Skoumal, Prince Birring, Michael Arango

We chose to use the Leaf Classification data set from kaggle (https://www.kaggle.com/c/leaf-classification/data) to train a CNN for classifying leaf specimens into one of 99 species of leaves. This data set contains 1,584 images of 16 samples of each of the 99 types of leaves. This data set should be large enough to sufficiently train a CNN to extract features in the leaves to be able to classify them. Our proposed plan is to treat this project as a research endeavor, ultimately resulting in developing a couple of different network architectures, we would also like to try to use a few different deep learning frameworks. The goal of our research project is to compare and contrast the pros and cons of the different network configurations and methods of implementation. We will split our data into training and testing sets to allow us to use a simple classification accuracy as our metric. If we have time, we would also like to compare our CNN network with some shallow learning classification models such as sklearn classifiers.