

COE 379L: Project 4 Proposal

Ashton Cole

April 15, 2024

1 Introduction

The goal of this project will be investigate how the architecture of a neural network impacts its accuracy. Neural networks fundamentally permit much leeway in their design. In a dense Artificial Neural Network (ANN), each layer can have an arbitrary number of perceptrons with a particular activation function. Compounding this, an arbitrary number of layers can be added. Each layer adds a whole dimension of complexity, and the possibilities are greatly increased when other types of layers, like convolutional, pooling, and dropout ones, are brought into consideration. The large *hyperparameter* space makes it intimidating to fine-tune a neural network model.

This project will conduct a parameter sweep of dense ANN architectures to glean any apparent patterns. This will hopefully identify both the optimal architectures and relationships between hyperparameters and accuracy.

2 Dataset

Two datasets will be investigated for this study. This will help evaluate whether or not results are tied to a particular dataset.

First, the hurricane damage dataset from the prior project will be revisited. This set contains 17,057 128×128 -pixel color satellite images of properties labeled as either damaged or not damaged by a hurricane.

In addition, a new *beans* dataset from the TensorFlow datasets module will be considered. This set contains 1,295 500×500 -pixel color images of leaves of beanstalks. Leaves are either categorized as healthy or having bean rust or angular leaf spot disease.

3 Techniques

The Keras Tuner module will be used to conduct a sweep of dense ANN architecture parameters. To simplify the sweep, activation functions will be held constant: the hidden layers will use the ReLU function, and the output layer will use the softmax function. Only the number of layers and the numbers of perceptrons they contain will be varied. Further, the number of layers used in each model will likely be limited to three, because additional layers increase complexity exponentially. They also increase the chances of encountering the vanishing gradient problem, which arises from back propagation, i.e. a “long” derivative chain rule when attempting to minimize the cost function.

Accuracy will be used as a simple metric to evaluate the quality of models. Testing accuracy will also be compared against training accuracy, to ensure that overfitting is not an issue. Additional metrics will be reviewed for the optimal models for each data set for further insights.

4 Deliverables

Since the main goal of this project is to study the networks themselves, the primary deliverable will be the data and analysis in the report. Jupyter Notebook files of the parameter sweeps will also be included for reproducibility.