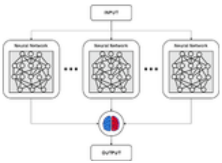
Uncertainty Quantification in Ensemble Deep Learning

Emily Diegel, Rhiannon Hicks, Max Prilutsky, and Rachel Swan
Mihhail Berezovski Ph.D, Jesse James Ph.D
Embry-Riddle Aeronautical University (ERAU)
Nevada National Security Site (NNSS)

Abstract

Data science is a broad field that encompasses studying and extracting meaning from data, and a subset of this category is machine learning. Machine learning is the tactic of using various techniques to build computer models that are able to learn by themselves using data. A neural network is a type of machine learning that is modeled after how the human brain learns, and takes a given input and predicts an output from this value. Neural networks are extremely applicable in the field of data science, but also come with one large drawback: no method exists for quantifying the amount of error in a neural networks output prediction. This begs the question of what methods we can attempt to implement or create to effectively measure the uncertainty in neural networks so that we can better comprehend and dissect the meaning of their predictions. The Research Experience for Undergraduates program students will attempt to create an ensemble neural network with a valid uncertainty quantification method. Throughout this paper we will explore the complete method that the REU program has attempted in order to combat this problem.



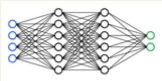
Nevada National Security Site

The Nevada National Security Site (NNSS) is a United States Department of Energy reservation located approximately 65 miles northwest of Las Vegas. NNSS specializes in defense nuclear nonproliferation and the pursuit of science. They cannot conduct full-scale nuclear tests, but instead conduct controlled underground experiments and computer simulations to verify that the arms are safe and usable.

INTRODUCTION

Neural networks provide the opportunity to analyze large amounts of image data to produce a regressive output. They are an algorithm modeled after the neurons in the human brain. Using a set of inputs and a set of desired outputs, the network is trained and tested in order to predict the desired outputs. The network trains itself through learning patterns and correlations between data and assigning weights and biases to certain values to produce a label.

Uncertainty Quantification (UQ) is the science of using various methods to quantify and reduce the amount of uncertainty in a data set. An ensemble neural network combines multiple similar but individual neural networks to decide on the most likely cumulative output based on each networks individual results. An ensemble is a type of UQ approach because it adds more checks to the neural network to ensure it is as certain as possible about its output value.



DATASET

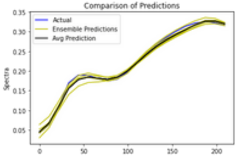
The dataset used is synthesis, optical imaging, and absorption spectroscopy data for 179072 metal oxides. The data contains 180902 images that are size 64 by 64, with 3 channel (RGB). Some of the images are blank, leading to 179072 total samples. The data is normalized from 0-1 for every channel. The input is a metal oxide where the color in the middle of each image represents the color of the printed sample, while the "coffee ring" around each sample is darker due to drying. The output is a spectra.

ARCHITECTURE

The layers for the neural network are as follows: Convolutional, Dense, Max Pooling, Flatten, Dense, Dense, Dense. The ensemble neural network was trained n = 10 times with 32 batch size, early stop epochs, and 40,000 images.

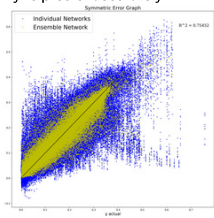
Results

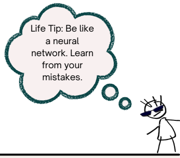
The figure gives an example output of the neural network, displaying the ensemble neural network prediction in comparison to the individual network predictions. The ensemble takes an average of each of the individual neural networks to produce a single output. Time constraints only allowed us to build an ensemble that averages all the network predictions equally, but a way to improve upon this would be to weight networks that had a lower mean squared error more significantly than networks with a high mean squared error.



Analysis

This graph represents the distribution of all predicted points on the spectra graphs. It is obvious that the ensemble does increase the accuracy of the predicted points since the yellow cloud is much more condensed around y = x than the blue cloud of points. Since in a perfect model the R^2 value would be equal to 1.00 (where all the predictions were 100% accurate), the figure shows that a decent R^2 value of 0.75452 was obtained, but it is clear that there are some limitations to our models ability to predict accurately.





Support for the program has been provided by the National Science Foundation (NSF) through REU Award Number DMS - 2050754.