Final Project Proposal

Kacie Salmon

Dataset

The dataset that I will be using for the final project is the Fashion MNIST benchmark dataset. This dataset consists of different clothing items to be classified. This dataset consists of 70,000 grayscale images of these clothing items, each a 28 x 28-pixel square made up of 784 pixels. Each pixel in these 784 are located through a number assigned based on the row and column. Each pixel in the dataset has a pixel-value, an integer from 0 to 255, that indicates lightness or darkness with 255 being the darkest. And each image also has an assigned label, an integer from 0 to 9, indicating what kind of clothing item it is. These images and the associated data are also broken into training and test data that can be used to train and test classifier algorithms on learning and applying correct labels to these images from the data given. There are 60,000 training images and 10,000 test images. The dataset is sourced from https://www.kaggle.com/zalando-research/fashionmnist.

Algorithms

The three algorithms I will be applying to the dataset are **least squares**, **support vectors**, and **neural networks**. For these algorithms, I will also experiment with data preprocessing and regularization if it improves the speed or optimization of the algorithms. For neural networks, I will experiment with the number of layers appropriate for the classification of this problem. To validate these algorithms and compare their performance, I will perform cross-validation with the given dataset and compare the overall average cross validation error for each algorithm, as defined from the misclassifications for each validation test. I will do seven validation tests for the cross-validation of each algorithm in which I hold out a group of 10,000 images.

Project GitHub

https://klsalmon.github.io/ECE532_FinalProject/

Timeline

Task	Timeframe
Application of Least Square	Oct 22 - Oct 29
Application of K-Means Clustering	Oct 29 - Nov 5
Application of Neural Networks	Nov 5 - Nov 17
Comparison of Validation	Nov 17 - Nov 24
Writing Final Report	Nov 24 - Dec 12