

# Machine Learning Final Project Proposal

For our project, we will use a relatively simple classification problem to compare neural networks and decision trees performance, train time, and scores.

## Group

This project will be completed by Ian Shakespeare and Dylan Sharp.

## Dataset

We will be using student performance data provided by University of California Irvine. The data can be found here: (<https://archive.ics.uci.edu/ml/datasets/student+performance>). We will specifically use the math class performance data as it is a more widely applicable subject than portugese, meaning more diversity in the dataset.

## Goal

Decision trees and neural networks are two very different methods of classification. We want to see if there is a clear advantage to using one or the other when classifying simple datasets. We will define one as advantageous if it results in faster train times or higher f1 score.

## Method

Modifying the code supplied from previous assignments, we will perform a grid search on varying sizes of decision trees and neural networks. Time to complete these grid searches will be recorded as the performance for that specific learning model. After an ideal set of parameters is found, we will compare the test scores for our final correctness comparison.

## Possible Enhancements

Should this process be quick and/or otherwise too simple, we could also compare the results of a SVC. This would mean performing the same tests and comparing SVC results with that of the neural network and decision tree. Furthermore, we could also perform a comparison when working on complex data, though this would require another dataset.