My data consisted of taking non-structured data or data that wasn’t very quantitative and assigning values to the data. It used audio frequencies and somehow figured out how to take measurements from the spectrograms.

A white background with black text

Description automatically generated

The testing error rate of the Boosting model was 0, which was encouraging. It seemed like only the Pruned decision tree and the Extreme Boosted model had a noticeable decrease in performance, but all models were under 0.5% error rates. I am interested as to why XGBoost did not perform as well. It didn’t have an accuracy output as the other models did in the cv step, therefore I think this may have affected the output of the error rate. There were many other parameters that the help documentation went over a bit, but honestly, I didn’t understand the affect of one versus the other, so likely something like eta or learning rate could have been impacting the results in a way I don’t understand yet.