



Rensselaer

why not change the world?®

Lab 3: Classification & Clustering with Abalone

Ahmed Eleish

ITWS-4600/ITWS-6600/MATP-4450/CSCI-4960

October 3rd, 2025

Tetherless World Constellation
Rensselaer Polytechnic Institute



Lab 3



<https://rpi.box.com/s/3be3ecqmxxyof1z8jv6f58tjfb89stl80>

Exercise 1:

- Train and evaluate 2 kNN models using the abalone dataset
- Each model should use a different subset of features
- Compare models using contingency tables
- Find the optimal value k for the better performing model by training over a range of k values and choosing the value with the highest accuracy (accuracy = correct classifications / total observations)

Exercise 2:

- Train K-Means and PAM models for the abalone dataset using the best performing feature subset from the previous exercise
- Find the optimal value K for each model and plot a silhouette plot of the model with that K value.

Please push to your github repository:

1. All your code in a *.R or *.MD file
2. A doc(x)/pdf of the text outputs + figures.

Thanks!
Have a great weekend!*