Lab 3: Classification & Clustering with Abalone

Ahmed Eleish
ITWS-4600/ITWS-6600/MATP-4450/CSCI-4960
February 13th, 2026

Lab 3

Files: https://rpi.box.com/s/5mavg7bl9pin6vrfy2l6ks5y81emdutd

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 from above by training over a range of k values and choosing the value with the highest accuracy (accuracy = correct classifications / total observations)

Exercise 2:

- Use k-means and PAM to find clusters in the abalone dataset using the feature subset from the exercise 1.

- Find the optimal value k for each model and plot silhouette plots for both models with their optimum k.

Please push to your github repository:

1. All your code in a *.R or *.MD file

Thanks!
Have a great weekend!*