## CS6140, Fall 2021

Mid-term Exam

Due: 10/24/2021, 11:59pm

1. Using the [Abalone Data Set](http://archive.ics.uci.edu/ml/datasets/Abalone), build a system to predict the age of an abalone from physical features. Try both a multi-class classifier and a regressor. For the regressor, you decide what a correct output is (within x from truth). The description of the data is [here](http://archive.ics.uci.edu/ml/machine-learning-databases/abalone/abalone.names). The dataset is [here](http://archive.ics.uci.edu/ml/machine-learning-databases/abalone/abalone.data).
2. Using the [Rain in Australia](https://www.kaggle.com/jsphyg/weather-dataset-rattle-package/version/1) data set, build a system that predicts whether it is going to rain tomorrow. The labels are Column W in the CSV file, so do not use Column W in your model.
3. Using the [Boston data set](https://www.cs.toronto.edu/~delve/data/boston/bostonDetail.html), build a linear regressor that predicts NOX and another one that predicts median home value.

Be good scientists, and don’t make decisions solely to maximize your results. Make decisions that correctly model the problem. You must use both logistic regression and gradient-boosted trees, and you must use train/test split. Don’t forget about regularization 😊.

To be turned in: Three distinct Jupyter notebooks with the code for the systems you build *as well as* write-ups describing the problem, the data, the decisions you made, the system choices you made and the results.