

COMP4432 - Assignment 1

(due by midnight MST the day prior to Live Session 2)

- **Assignment 1, Part 1:** Data Loading and Preparation. Load the diabetes dataset into two numpy arrays: one for the feature set and one for the target. Pick a single feature to try to predict the target (disease progression). Document the reason you chose the feature you did. Break your single feature and target sets into training and test sets with the last 20 rows being in the test set.
- **Assignment 1, Part 2:** Model Training. Instantiate a linear regression model, and train it with your single feature and target sets.
- **Assignment 1, Part 3:** Prediction and Measurement. List the first 10 predictions on your single feature training set. Print out the feature coefficient and the root mean squared error of your model. [use test dataset](#)
- **Assignment 1, Part 4:** Visualization. Print out a scatter plot with the feature you chose on the x-axis, and progression on the y-axis. Plot the regression line on this same graph with appropriate labels on each axis.