**Regression Improvements and Random Forests**

Walkthrough Video:

<https://berkeley.zoom.us/rec/share/aCKX8Zsgf6c3SEZYAThF-6DuRf6Go5Irm5q85YIaUNYSAkkEDu4T-NRFarCpRFee.B15Qu29S3mgJiJrB>

This is going to be the last week of new content for this project. Since it is relatively dense and more than the previous weeks, I would recommend spending a bit more time on it to really make sure you can understand the concepts well. For this week, we explored a few minor improvements towards our linear regression model from last week, and then introduced a new technique called random forest regression to build a better model predicting charges. Here are some general things for you to consider and try to answer on your own:

* Where do our scores come from? Understand this and how you would calculate it if you were forced to do it by hand.
* Why do we conduct Cross Validation?
* Make sure you are comfortable with how decision trees in general work. Feel free to look around online for other resources. The link to the manual of the two models we used in the walkthrough are here: [Linear Regression](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html), [Random Forest Regressor](http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html)

When it comes to building off of this week’s content, there are a few different approaches you can do this with in varying degrees of difficulty.

* The simplest thing is to try this out with different quantitative data outside of just charges, or to tinker around by dropping more variables from the classifier to see if you can achieve a better accuracy.
* Try to research different models! (Lasso, Polynomial, Ridge)
* Most of what has been done is on regression. Classifiers are the counterpart meant for qualitative data, so you can try and conduct random forest classification to predict something such as whether a person is a smoker or not.
* As always, feel free to try anything you have in mind as well.

Again, a lot of the heavier context behind the things shown in this week’s content has been omitted without deep explanations. The bulk of this project has been designed with the focus of a “hands on” experience to try out key techniques rather than go deep into the theory behind things, but it can be very helpful to spend time on your own to try and learn more about what was shown. It can also help you be more comfortable with designing unique things that make this project your own.