

# HW 7 - Neural Networks

*William Morgan, Jared Scolaro, Mitch O'Brien*

*April 9, 2018*

---

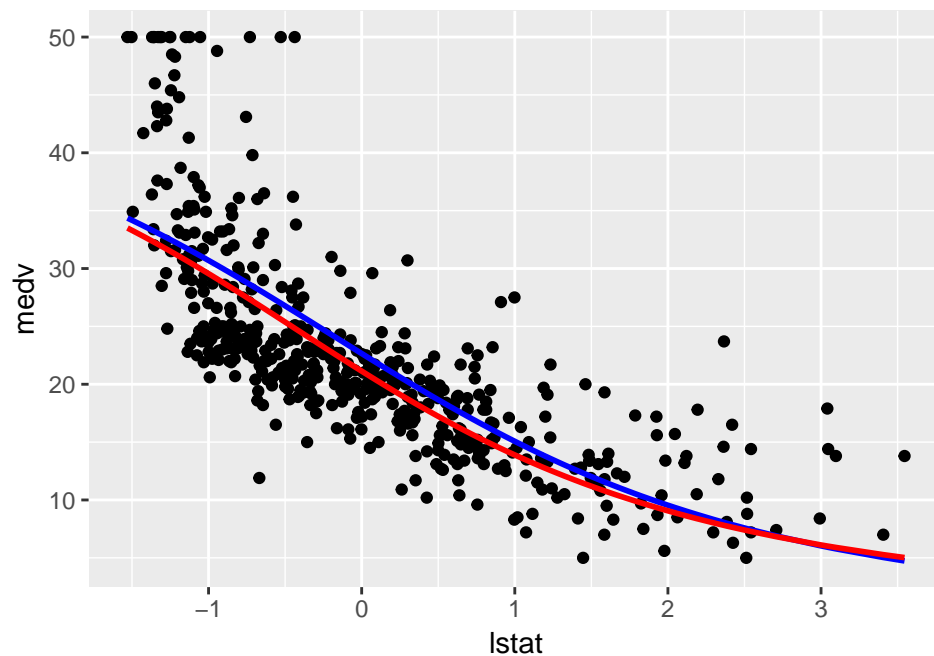
## 1. Boston Housing with a Single Layer

- Fit the model with 100 units, decay = .001, and plot the fits. How does it look? Try running the fit at least twice to see that it changes

```
model <- keras_model_sequential()

model %>%
  layer_dense(units = 100,                # input layer
              activation = 'sigmoid',
              input_shape = c(1),
              use_bias = TRUE) %>%
  layer_activity_regularization(l2 = .001) %>% # regularization on input
  layer_dense(units = 1)                  # output layer

model %>%
  compile(
    loss = 'mean_squared_error',
    optimizer = optimizer_sgd())
```



The two fits are pretty similar, as expected. Furthermore, both of them appear to be pretty simplistic in the way they describe the relationship between `lstat` and `medv`. It might be reasonable to reduce the amount of regularization (i.e. increase complexity) we're adding in order to capture more non-linearity.

- Redo the loop over decay values with 100 units. How does it look now? Do we need 100? Will decay be more important with 100 than it was with 5 units?

We should expect to see regularization have less of an effect on the model containing only 5 units, as there are significantly many more predictors in the network with 100 units. The regularization will thus have a larger effect on the absolute value of each individual parameter in the 100-unit model compared to the 5-unit model.

TODO: plot results of looping through decays