

STAT 6240 - HW 5 - Due 4/04

Due Date: April 4th, Thursday, 6:10 PM

Your homework submission should contain:

1. Your R code,
2. Outputs from R (including figures),
3. Answers to the questions.

Do not print out your homework. Save it as a pdf file and upload it to Blackboard.

1. Read Chapter 3 of Deep Learning with R (available at <https://livebook.manning.com/#!/book/deep-learning-with-r/chapter-3>). The chapter also has the R code we used for the Reuters and Boston datasets.
2. Setup Tensorflow and Keras on your own computer See <https://keras.rstudio.com/> for instructions.
3. Load the Boston Housing dataset from the MASS package.
4. Read the dataset description (`?Boston`) and identify which variables in the dataset are categorical variables (*Hint:* there are two). Use the `factor` and `model.matrix` functions to create new dummy variables for these categories. This is referred to as “one-hot coding” in the book.
5. Use the `caret` package to tune the parameters of a one-hidden-layer neural network. You can find instructions for `caret` at <http://topepo.github.io/caret/index.html>. Also, see the `caret_example.R` file for an example. You will perform 10 fold cross-validation on the dataset and compare the “mse” of each model to find the ideal parameters for number of nodes in the hidden layer (`size`), the dropout rate (`dropout`), the training batch size (`batch_size`), the learning rate (`lr`) and the activation function (`activation`, possible choices are `relu`, `sigmoid` and `tanh`). Instead of comparing all possible choices at once, you should fix some of the parameter values and optimize over the others. That is, instead of supplying a `tuneGrid` with all possible values of the parameters, use a grid in which only two variables’ values change, and fix other values to a reasonable number.
6. Next, fit the model with the best parameters using Keras. Obtain its predictions. To make sense of the results, plot the predictions with respect to each variable. Summarize your findings. Which variables seem to have non-linear effects? Which variables are significant? **In class, we will discuss this part in detail.**