**Week 9 Paper**

Kyle Parker

Here is what I took away from the chapter of the “Practical Data Science with R” reading:

1. Our main job as data scientists is to solve real world problems. This is done by choosing a method for a given goal, and applying a scientific model. This model then needs to be evaluated on training data as well as testing data, to ensure quality of the model and to make sure that it will work in production as it worked in testing.
2. Mapping problems to certain machine learning tasks is a very important part of solving a problem. These different machine learning tasks (scoring, classification, problems without a known target) all work very differently. For certain applications, one of these tasks will be much better suited than others, and for other applications, certain tasks may not work at all!
3. Even after choosing an appropriate machine learning task for your problem, evaluating and validating the model you have created is very important. The model must be evaluated in order to ensure that it is optimal, and then must be validated in order to ensure that it will work in production and on “real world” data as well as it did with the test data.