**Chapter 2: My first model**

This chapter will feel for some of you as Hello World to machine learning, and that's how it needs to feel. We will take a simple model start (and that is why it can be said by a beautiful analytical thing that cannot be said for more complex models such as elegance) and try to use it on a real problem.

This exercise of running a model in reality inspires a lot of questions, how to examine the model, how to compare different models, while ignorance awaits me--a man of the beginning.

**This chapter will deal with regression problems – which may be the more intuitive side to define loss** **and understanding the framework** **of the standard.**

**Theoretical part:**

1. Read chapter 3 in ISLR - Linear Regression.
2. Read chapter 5 in ISLR - Resampling Methods.
3. Read chapter 6, parts 6.1 and 6.2 in ISLR - Linear Model Selection.

**Practical part:**

In the community of the data-scientist there are several known datasets which comprise a benchmark for the community. The data on the neighborhood prices in Boston is one of them.

The reservoir contains not many records, on neighborhoods in Boston and the average housing prices. Each neighborhood has several information items that can help predict the average price in the neighborhood.

The repository is among the first datasets that were created in the field, it is not large and therefore the question of the fitting rises more sharply.

Use sklearn load the data into your notebook.

**Then follow the steps of the classical study (of course it's just a template, and you should be making frequent iterations**  **with the tutor at this time, and repeat the steps**  **after the new comments):**

1. Perform Aksploratia: for example:
   1. Test of distributions and correlations
   2. Make sure you understand the meaning of the columns and the values that they contain
2. Clear the information: for example-
   1. Complete missing values
   2. Remove information which constitute anomalies (this step is not about running anomaly detection algorithms).
3. Run the model: for example-
   1. Constructing a model for all the information.
   2. Examination of the resulting model (weights, etc.)
   3. Examining errors or variances that are greater than the desired value.
   4. Training exam with and with existing features.
   5. Extracting significant features and a renewed training to improve the score.