

Assignment

Practice prediction with Deep Learning

Dr. Eran Raviv

July 2022

Submission date: 22nd of July, 23:59.

In this assignment you will “get your hands dirty”. You are given a dataset of house prices alongside variables which are supposed to explain it (e.g. size of the house, parking included yes/no). You need to predict the house price based on the set of the provided explanatory variables. You will use a deep learning model to do this.

The cost function used for evaluation is the Root Mean Squared Logarithmic Error. Why? the nature of the data is such the minimum price is capped from below, while the maximum is not. Some houses are very expensive, and a big error around those samples should not dominate the cost function. Specifically, if N is the number of (test) samples, \hat{y} and y are the predicted and actual house prices, then the evaluation metric is:

$$\sqrt{\frac{1}{N} \sum_{i=1}^N (\log(\hat{y}_i) - \log(y_i))^2}$$

Detailed Instructions

- (a) You are given a benchmark model.
- (b) Create your own (deep learning) model, or change/extend the benchmark model.
- (c) Budget few observations for validation of your model.
- (d) Once you are comfortable with your choices, you can use the benchmark script which is given to create your vector in a CSV format for submission. The CSV file name should be named using the following format: "surname.csv" where *surname* is your surname.
- (e) Submit via this webapp. You can submit only once. Your accuracy score is directly reported on the screen.
- (f) While accuracy has it's place, the goal is to gain hands-on experience. Write a report outlining your efforts. Provide your own opinions. For example:
 - What is the architecture you eventually settled on (layers..., neurons..)? why?
 - What hyperparameters did you find important in your experiments?
 - How did you estimate the expected error of your model?

Report length: max one page (you can place tables or charts in an Appendix); file-format should be pdf.

Good luck and have fun!