

Deep Learning for NLP

Younes Samih

Department of Computational Linguistics

November 22, 2017

Road Map

Boston Housing

Boston Housing Dataset

Boston Housing

- ▶ House prices prediction using neural Network
- ▶ Regression Problem

Boston Housing Dataset

- ▶ The data¹ contains 14 columns
 - ▶ **CRIM**: per capita crime rate by town.
 - ▶ **ZN**: proportion of residential land zoned for lots over 25,000 sq.ft.
 - ▶ **INDUS**: proportion of non-retail business acres per town.
 - ▶ **CHAS**: Charles River dummy variable (= 1 if tract bounds river; 0 otherwise).
 - ▶ **NOX**: nitric oxides concentration (parts per 10 million).
 - ▶ **RM**: average number of rooms per dwelling.
 - ▶ **AGE**: proportion of owner-occupied units built prior to 1940.

¹[https://](https://archive.ics.uci.edu/ml/machine-learning-databases/housing/)

archive.ics.uci.edu/ml/machine-learning-databases/housing/

Boston Housing Dataset

- ▶ **DIS**: weighted distances to five Boston employment centers.
- ▶ **RAD**: index of accessibility to radial highways.
- ▶ **TAX**: full-value property-tax rate per \$10,000.
- ▶ **PTRATIO**: pupil-teacher ratio by town.
- ▶ **B**: $1000(Bk - 0.63)^2$ where Bk is the proportion of blacks by town.
- ▶ **LSTAT**: % lower status of the population.
- ▶ **MEDV**: Median value of owner-occupied homes in \$1000s.