L02

Linear Regression
Different Approaches

Demo

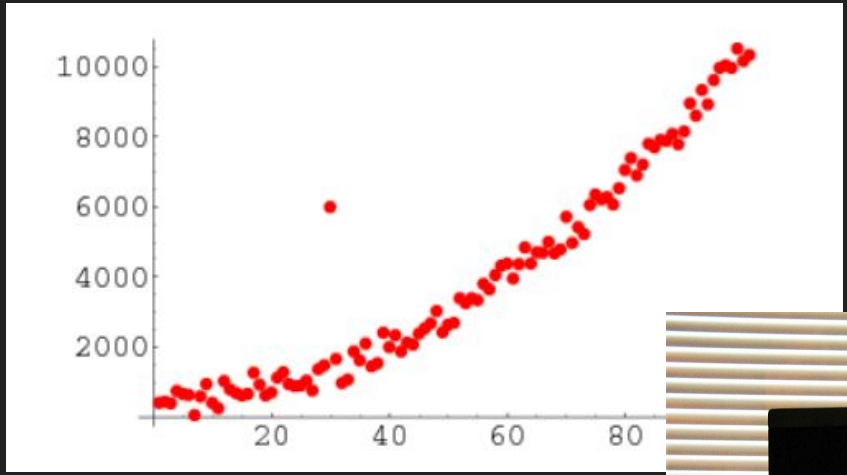
Demo

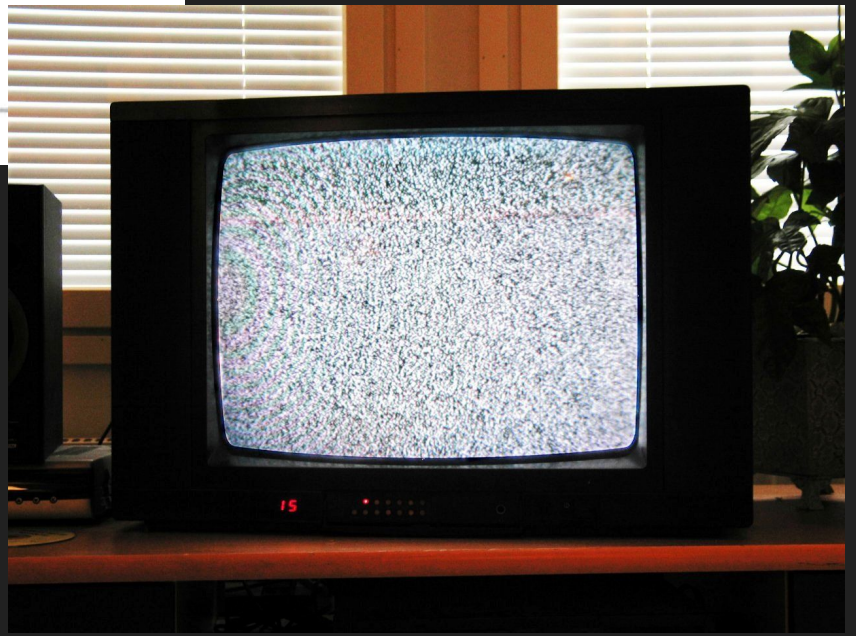
Your regular routine:

```
$ python -m venv venv
$ source venv/bin/activate
$ pip install -r requirements.txt
```

The Data

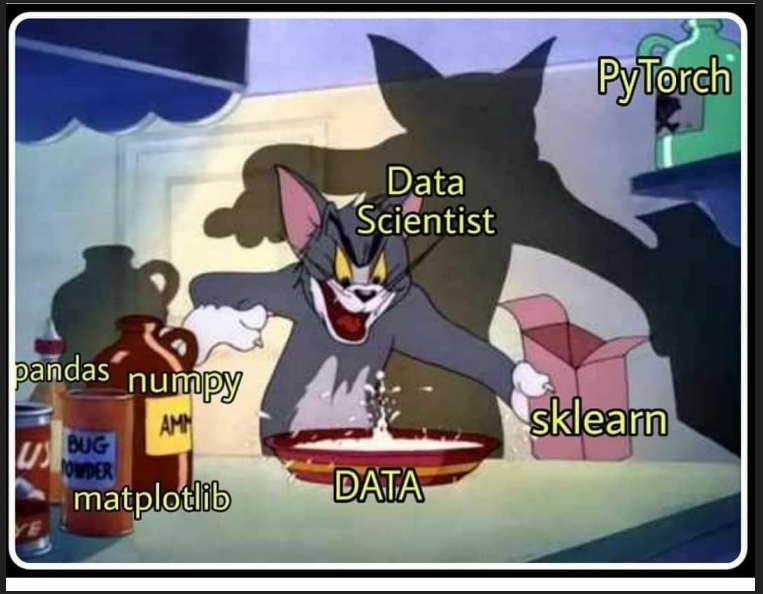
- Data
- Noise
- Outliers





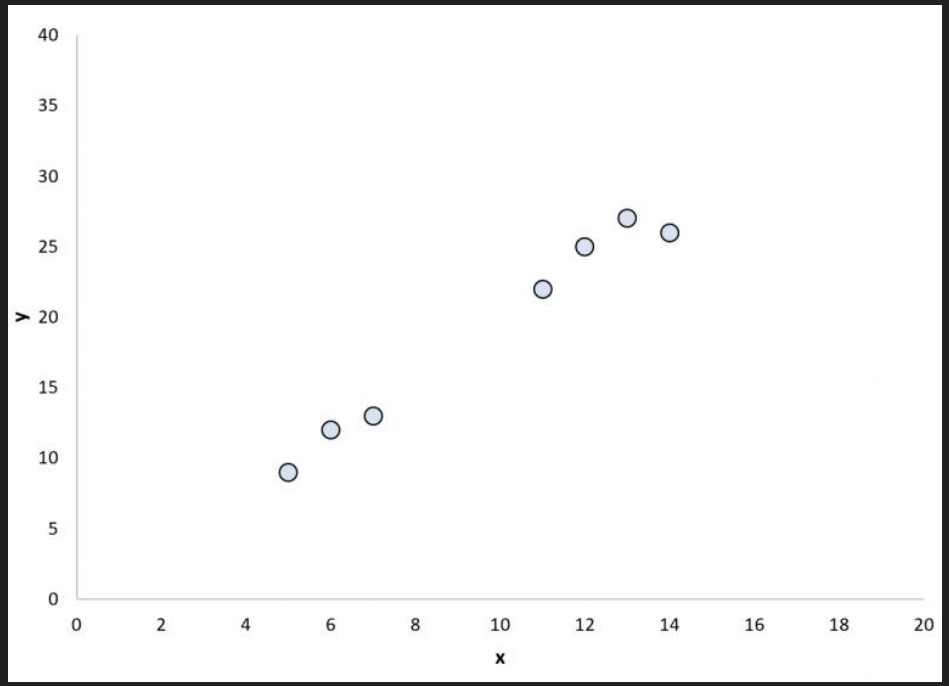
Project Stages (Simplified)

- Visualization
- Exploratory data analysis
- Outlier detection, data cleaning
- Feature selection and engineering
- Optimization function formulation
- Experimenting with architecture, hyperparameters, optimization algorithms

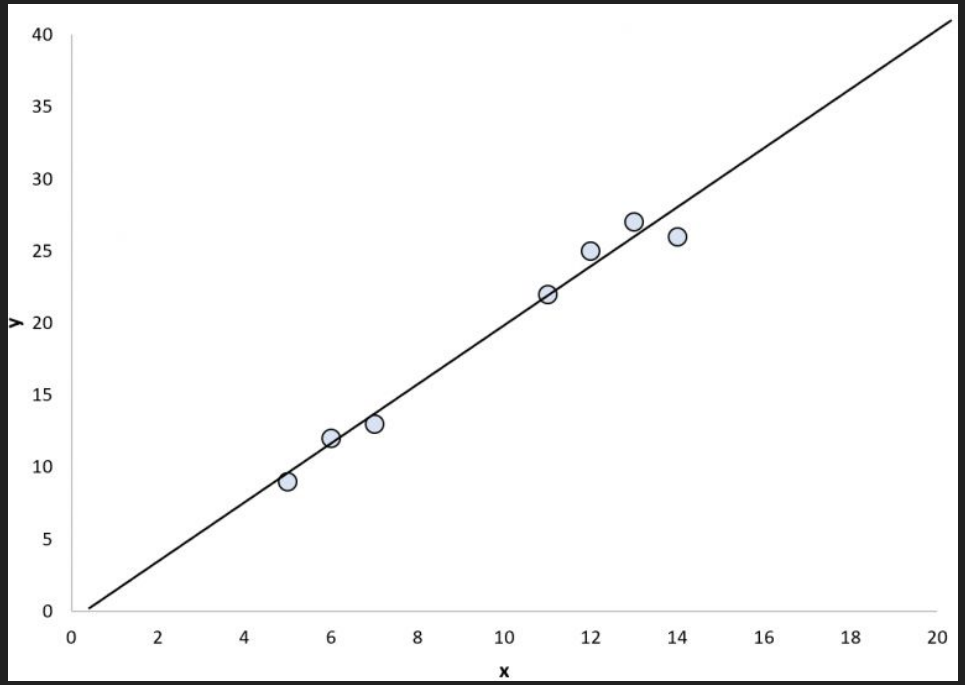


TODO: show a simple project (wine dataset)

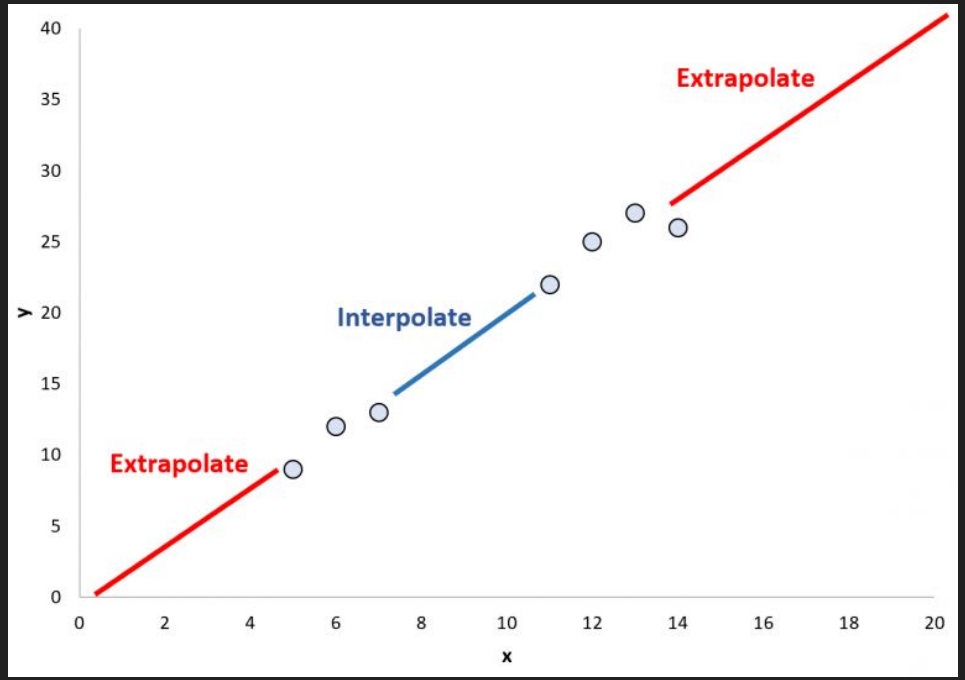
Is Our Model Ok? Interpolation vs. Extrapolation



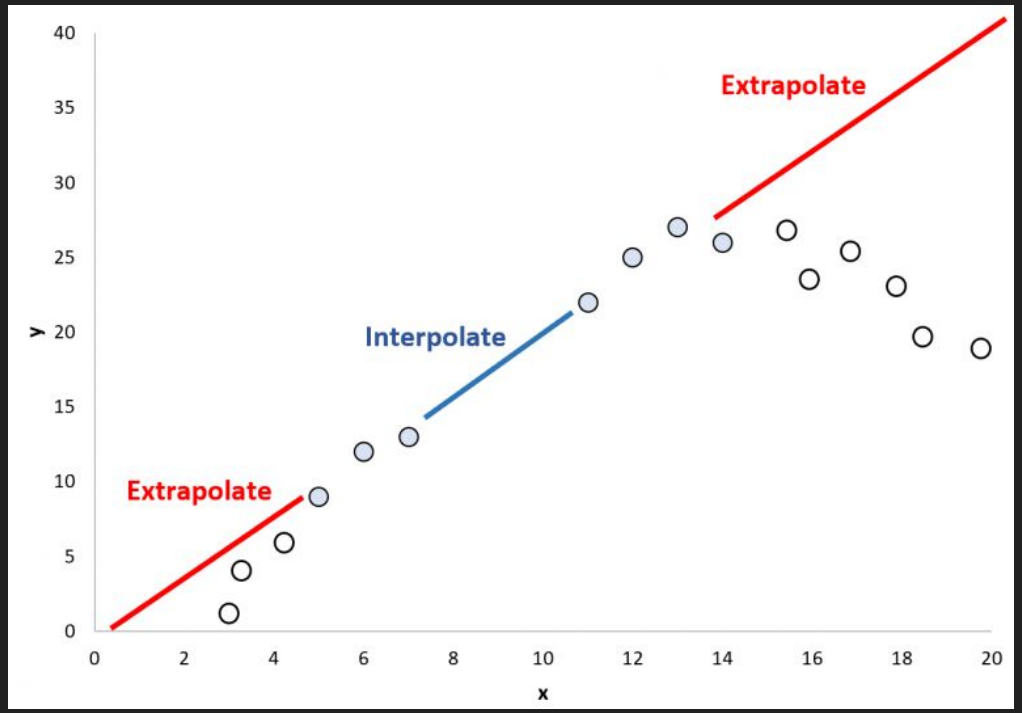
Is Our Model Ok? Interpolation vs. Extrapolation



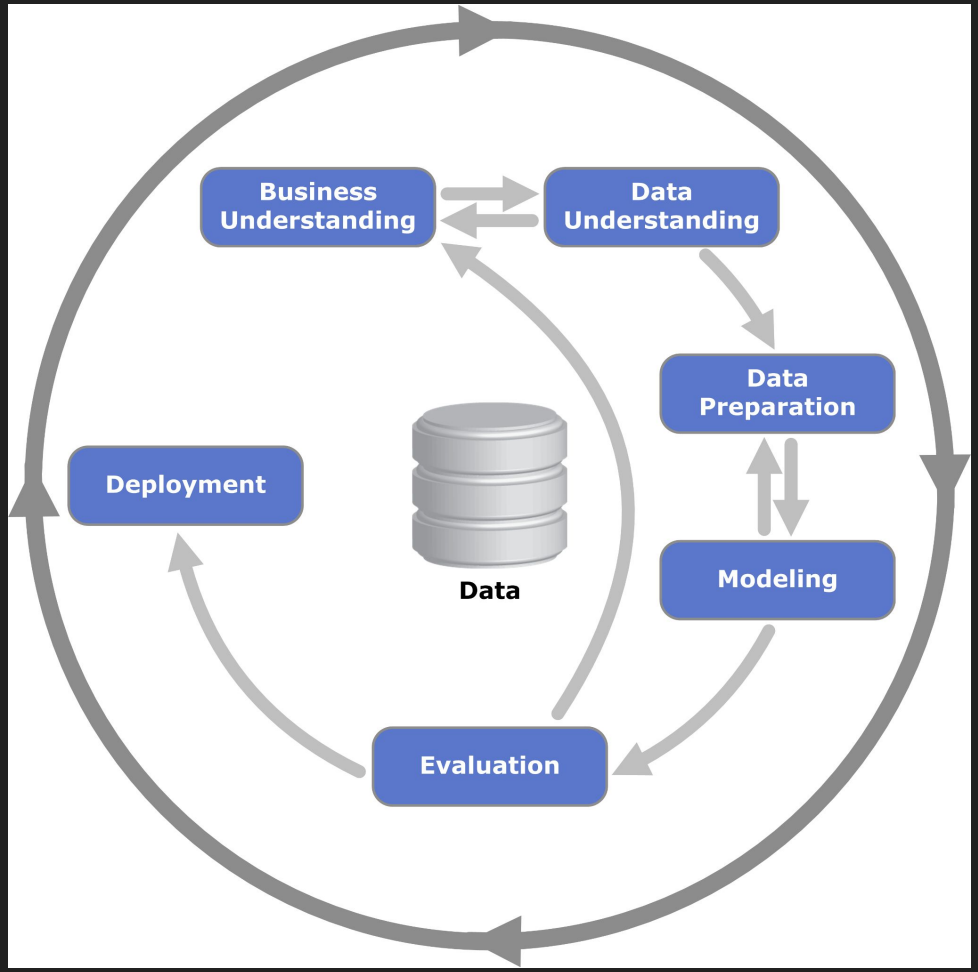
Is Our Model Ok? Interpolation vs. Extrapolation



Is Our Model Ok? Interpolation vs. Extrapolation



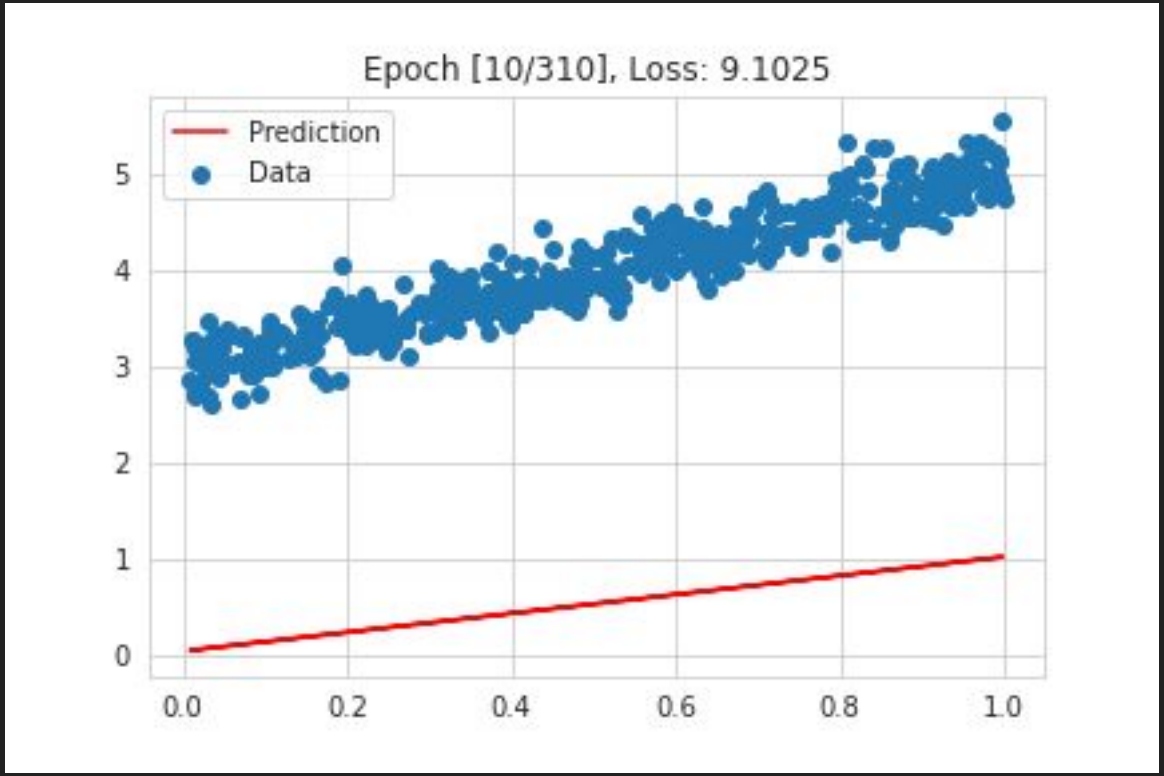
CRISP-DM


Business Understanding
Data Understanding
Data Preparation
Modeling
Evaluation
Deployment
Data

HW

- Plot loss function value (should drop over the fitting, loss = f(epoch))
- Try RMSE, MAE and maybe other losses for linear regression
- Make animation for fitting: plots of changing fitting curve (line) over the data (see next slide)
- Experiment with non-linear data, for example:
  - y = 2 * x**2 + x + 3.5 + noise
- Experiment with number of samples, sigma, and optimization algorithms

HW


Epoch [10/310], Loss: 9.1025
Prediction
Data
5
4
3
2
1
0
0.0
0.2
0.4
0.6
0.8
1.0

Some References

Regression in matrix form:

- https://online.stat.psu.edu/stat462/node/132/

OLS results explained:

- https://medium.com/swlh/interpreting-linear-regression-through-statsmodels-summary-4796d359035a
- https://towardsdatascience.com/simple-explanation-of-statsmodel-linear-regression-model-summary-35961919868b

Statsmodels docs:

- https://www.statsmodels.org/stable/examples/index.html

Some References

- Probability vs Likelihood : https://sebastianraschka.com/faq/docs/probability-vs-likelihood.html