Machine Learning

Practical Sheet 2: Linear Regression

Consider the "Pizza" dataset, provided as a ".csv" file in the course web page. We are interested in predict the "Startup Cost" value, supposing that it is only dependent of the "Annual Fee" feature. Of course, this is an over-simplification, just for "toy-problem" purposes.

Annual Fee	Startup Cost
1000	1050
1125	1150
1087	1213
1070	1275
1100	1300
1150	1300
1250	1400
1150	1400
1100	1250
1350	1830
1275	1350
1375	1450
1175	1300
1200	1300
1175	1275
1300	1375

1. Implement one Python script to generate, by brute-force, "N" hypotheses of linear regression models, and keep the best one.

2. Implement one Python script to obtain the closed-form solution of a linear regression problem with a single independent variable.

3. Implement one Python script to obtain the gradient descent solution of a linear regression problem.

   a. Using a fixed number of iterations as stop criterium;

   b. Stopping the algorithm when no relevant improvements occur;

4. Implement a computationally more efficient version of linear regression: batch linear regression.

5. Design and implement an "adaptive learning rate" strategy, where the idea is to use larger steps at the beginning of the iterative process, and iteratively decrease them once the solution is close.

6. Compare the computational/accuracy performance of the fixed vs adaptive learning rate strategies.