

MAI391 – ASSIGNMENT 3

Deadline: 23h50', 14/07/2024

In this assignment, you are required to work in a group and complete the given tasks. Namely, you will implement **gradient descent** and **linear regression** from scratch using Python with the help of following tasks:

1. Look for a suitable secondary data

Hint: *Any topic of your choice with a minimum of 100 observations.*

2. Import Libraries and load the data

Hint: *Import essential modules and helper functions from NumPy and Matplotlib.*

3. Visualize the Data

Hint: *To understand the data, you can use a scatter plot for visualization.*

4. Compute the Cost Function $J(\theta)$

Hint: *The objective of linear regression is to minimize the cost function*

$$J(\theta) = \frac{1}{2n} \sum_{i=1}^n (f_{\theta}(x^{(i)}) - y^{(i)})^2$$

where $f_{\theta}(x)$ is given by the linear model $f_{\theta}(x) = \theta^{\top} x = \theta_0 + \theta_1 x_1$.

5. Implement Gradient Descent

Hint: *Minimize the cost function $J(\theta)$ by updating and repeat the below equation*

$$\theta_j := \theta_j - \gamma \frac{1}{n} \sum_{i=1}^n (f_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)}$$

(simultaneously update θ_j for all j).

6. Visualizing the Cost Function $J(\theta)$

Hint: *You will plot the cost over a 3-dimensional grid of θ_0 , θ_1 and J_{θ} values.*

7. Plotting the Convergence

Hint: *Plot $J(\theta)$ against the number of iterations of gradient descent.*

8. Training Data with Univariate Linear Regression Fit (Best Fit Line)

Hint: *Correctly implemented and run gradient descent and arrived at the final parameters of the model, we can use these parameters to plot the linear fit line.*

9. Inference using the optimized θ values

Hint: *Using the final values for θ to make predictions, $f_{\theta}(x) = \theta^{\top}x$.*

Submission: Each submission **must** consists of

- Your Python source codes.
- A brief report explaining your work.