Math Basics for Machine Learning
Graded Assignment 4

Your Name Here

Fall 2023

Instructions

This is the fourth graded assignment for the Math Basics for Machine Learning course. It contains two tasks. The instructions, as well as links to supplementary material, are given in the task descriptions.

Provide **detailed solutions** to the tasks in this assignment. Then, save your solution document as a .pdf file and submit it by filling in the corresponding Google form.

In total, you can earn 10 points for this assignment. This score will contribute to you final score for this course.

You must submit your answers by **Monday, November 6, 18:59 Moscow Time**.

Solutions must be typed in LaTeX. Hand-written solutions, as well as late submissions, will not be accepted.

It is the idea that you complete this assignment individually. Do not collaborate or copy answers of somebody else.

Have fun!

1. (4 points) Find and classify all the critical points of the following function

$$f(x, y) = 7x - 8y + 2xy - x^2 + y^3$$

Solution: Your solution here

2. (6 points) Fitting a machine learning model means finding the optimal values of its parameters, which comes down to optimizing some loss function $\mathcal{L}$. In class, we saw the least-squares example. Now, let's consider a so-called *logistic loss*:

$$\mathcal{L} = \sum_{i=1}^{n} \left[y_i \log \sigma_i + (1 - y_i) \log (1 - \sigma_i) \right],$$

where $\sigma_i = \sigma_i(w_0, w_1) = \frac{1}{1+\exp -(w_0 + w_1 x_i)}$.

Here, $\{x_i, y_i\}_{i=1,\ldots,n}$ are the observed data points, and w_0 and w_1 are the parameters of the model.

(a) (4 points) Find the gradient of the loss function above.

Solution: Your solution here

(b) (2 points) Suppose that you have a single observation:

$$x_1 = 2,\ y_1 = 1.$$

Let's assume the initial weights w_0 and w_1 are both set to 0. Perform one step of a gradient descent update of the parameter values. Use learning rate $\eta = 0.1$

Solution: Your solution here