COMPSCI 4ML3, Introduction to Machine Learning
Assignment 1, Fall 2024

Hassan Ashtiani, McMaster University

Due date: Monday, October 7, 11pm

This assignment has a programming component (i.e., question 1). You should upload two separate files as follows. Avoid compressing these files to make the marking process easier.

- One pdf files (call it MacID.pdf) containing your **typed** solutions. This file includes report for the programming component (answers to all questions including the results, plots, and discussions) as well as the solutions to other (theory) questions. The report for the programming component should be self-explanatory (TAs should understand the solutions without the need to consult the code)
- One Jupyter notebook file (call it MacID.ipynb) containing all your codes for question 1.

If you have questions about the assignment, please ask them in the Q/A channel.

---

1. [**70 points**] Programming Component. Follow this link to open the Google colab environment and make a copy of the notebook. Include the answers/graphs/pictures/analyses the tasks in your final pdf report. Additionally, upload your modified Jupyter notebook that includes your code (as a separate ipynb file).

2. [**15 points**] In this question we will use least squares to find the best line ($\hat{y} = ax + b$) that fits a non-linear function, namely $f(x) = 2x - x^3 - 1$. For this, assume that you are given a set of n training points $\{(x^i, y^i)\}_{i=1}^n$=$\{((i/n), 2(i/n) - (i/n)^3) - 1\}_{i=1}^n$. Find a line (i.e., $a, b \in \mathbb{R}$) that fits the training data the best when $n \to \infty$. Write down your calculations as well as the final values for a and b. (Additional notes: the $n \to \infty$ assumption basically means that we are dealing with an integral rather than a finite summation. If it makes it easier for you, instead of working with an actual training data you can assume x is uniformly distributed on $[0, 1]$.)

3. [**15 points**] In this question we would like to fit a line with zero y-intercept ($\hat{y} = ax$) to the curve $y = x^2$. However, instead of minimizing the sum of squares of errors, we want to minimize the following objective function:

$$\sum_i \left[\log\left(\frac{\hat{y}^i}{y^i}\right)\right]^2$$

   Assume that the distribution of x is uniform on $[2, 4]$. What is the optimal value for a? Show your work.