

CSCI 145 Problem Set 2

September 1, 2025

Submission Instructions

Please upload *your* work by **11:59pm Monday September 8, 2025**.

- You are encouraged to discuss ideas and work with your classmates. However, you **must acknowledge** your collaborators at the top of each solution on which you collaborated with others and you **must write** your solutions independently.
- Your solutions to theory questions must be written legibly, or typeset in LaTeX or markdown. If you would like to use LaTeX, you can import the source of this document (available from the course webpage) to Overleaf.
- I recommend that you write your solutions to coding questions in a Jupyter notebook using Google Colab.
- You should submit your solutions as a **single PDF** via the assignment on Gradescope.

Grading: The point of the problem set is for *you* to learn. To this end, I hope to disincentivize the use of LLMs by **not** grading your work for correctness. Instead, you will grade your own work by comparing it to my solutions. This self-grade is due the Friday *after* the problem set is due, also on Gradescope.

Problem 1: Single Value Functions

Consider a supervised learning problem with n labels $y^{(1)}, \dots, y^{(n)} \in \mathbb{R}$. In class, we explored the linear function class that predicted a weighted combination of the input points. In this problem, we'll consider the function class that outputs a single real number $m \in \mathbb{R}$ for all points.

A single number that best fits the data is known as its *central tendency* in statistics. Here, we will derive different central tendencies from an empirical risk minimization perspective.

Part A: ℓ_2 -norm

Consider the ℓ_2 -norm loss function

$$\mathcal{L}(m) = \sum_{i=1}^n (y^{(i)} - m)^2.$$

Show the optimal value m^* is the average.

Part B: ℓ_∞ -norm

Consider the ℓ_∞ -norm loss function

$$\mathcal{L}(m) = \max_{i \in \{1, \dots, n\}} |y^{(i)} - m|.$$

Derive the value m^* that minimizes this loss.

Hint: Think about the minimization problem directly rather than using derivatives.

Part C: ℓ_1 -norm

Consider the ℓ_1 -norm loss function

$$\mathcal{L}(m) = \sum_{i=1}^n |y^{(i)} - m|.$$

For simplicity, assume that n is odd. Show that the optimal value m^* is the median.

Hint: Try drawing the loss on top of a plot of the points on the number line.