

CS 4644/7643: Deep Learning

Fall 2026

Homework 0 Theory

Instructor: Danfei Xu

TAs: Rohan Bansal, Uzu (Chris) Lee, Alfred Cueva

Kun-Lin Hsieh, Chuye Zhang, Liqian Ma, Robert Azarcon

Dhruv Ahuja, Xinchun Yin

Discussions: <https://piazza.com/class/ms43b2tjxa63lq>

Due: Friday Aug 28 2026, 11:59pm ET

Instructions

1. We will be using Gradescope to collect your assignments. Please read the following instructions for submitting to Gradescope carefully! Failure to follow these instructions may result in parts of your assignment not being graded.
 - Submit your answers for Homework 0 Theory to the **HW0 Theory** assignment. For every question, there is only one correct answer. This will count for half of the points in HW0, which will be 0.5% of your total grade.
 - Submit your code to the **HW0 Code** assignment. Double-check that your submission is in the correct format and that the autograder runs successfully after your submission. This will count for half of the points in HW0, which will be 0.5% of your total grade.
 - Submit your notebook PDF to the **HW0 Notebook** assignment. This submission will not be graded, and will not count towards your final grade. It's purpose is for you to check that you can generate notebook PDFs correctly for future homeworks.
2. We generally encourage you to collaborate with other students. You may talk to a friend, discuss the questions and potential directions for solving them. However, you need to write your own solutions and code separately, and *not* as a group activity.
3. Homework 0 is intended as an exercise to check your knowledge of course prerequisites. If you find yourself having trouble completing these problems, you do not need to drop the class, but we would recommend brushing up on background in calculus, linear algebra, and statistics.

1. (1 point) Consider the tables below that display infection rates for a disease in two independent regions given vaccine status.

Region	Pop.	Vaccination Rates	% of Population Infected
Cityville	874,961	77.0%	0.36%
Townsland	578,759	37.7%	1%

Region	% of Infected people that are Vaccinated	% of Infected people that are Unvaccinated
Cityville	27.8%	72.2%
Townsland	5.0%	95.0%

It appears that infected individuals in Cityville are much more likely to be vaccinated than in Townsland. Given these tables, would a vaccinated individual be less likely to be infected in Cityville or Townsland?

☐ Cityville ☐ Townsland

2. (1 point)

Given a (possibly) biased coin with $P(\text{Heads}) = p$ and $P(\text{Tails}) = 1 - p$, first determine the method to generate a fair outcome (50:50) in the fewest amount of flips using this coin.

What is the expected number of coin flips required (in terms of p) to produce a fair outcome using this method?

☐ $\frac{1}{p(1-p)}$ ☐ $\frac{1}{1+p^2}$ ☐ $\frac{2p}{1-p}$ ☐ A fair outcome cannot be generated with a biased coin

3. (1 point) X is a continuous random variable with probability density function:

$$p(x) = \begin{cases} 2x^3/81 & 0 \leq x \leq 3 \\ 2(x-3)/8 & 3 < x \leq 5 \end{cases} \quad (1)$$

Which of the following statements are true about the equation for the corresponding cumulative density function (CDF) $C(x)$?

[Hint: Recall that CDF is defined as $C(x) = Pr(X \leq x)$.]

- ☐ $C(x) = x^4/162$ for $0 \leq x \leq 3$
☐ $C(x) = x^2/8 - 3x/4 + 13/8$ for $3 < x \leq 5$
☐ All of the above
☐ None of the above

4. (2 point) A random variable x in standard normal distribution has the following probability density:

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (2)$$

Evaluate the following integral:

$$\int_{-\infty}^{\infty} p(x)(ax^2 + bx - c)dx \quad (3)$$

[Hint: We are not sadistic (okay, we're a little sadistic, but not for this question). This is not a calculus question.]

- ☐ $a + b + c$ ☐ $-c$ ☐ $a - c$ ☐ $b + c$

5. (2 points) Consider the following function of $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6)$:

$$f(\mathbf{x}) = \sigma \left(\ln \left(5 \left(\max\{x_1, x_2\} \cdot \frac{x_3}{x_4} - (x_5 + x_6) \right) \right) + \frac{1}{2} \right) \quad (4)$$

where σ is the sigmoid function

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

Compute the gradient $\nabla_{\mathbf{x}} f(\cdot)$ and evaluate it at $\hat{\mathbf{x}} = (-1, 3, 4, 5, -5, 7)$.

- ☐ $\begin{bmatrix} 0 \\ 0.031 \\ 0.026 \\ -0.013 \\ -0.062 \\ -0.062 \end{bmatrix}$ ☐ $\begin{bmatrix} 0 \\ 0.157 \\ 0.131 \\ -0.065 \\ -0.314 \\ -0.314 \end{bmatrix}$ ☐ $\begin{bmatrix} 0 \\ 0.357 \\ 0.268 \\ -0.214 \\ -0.894 \\ -0.894 \end{bmatrix}$ ☐ $\begin{bmatrix} 0 \\ 0.357 \\ 0.268 \\ -0.214 \\ -0.447 \\ -0.447 \end{bmatrix}$

6. (2 points) Suppose you want to predict an unknown value $Y \in \mathbb{R}$, but you are only given a sequence of noisy observations $x_1, \dots, x_n$ of Y with i.i.d. noise ($x_i = Y + \epsilon_i$). If we assume the noise is I.I.D. Gaussian ($\epsilon_i \sim N(0, \sigma^2)$), the maximum likelihood estimate ($\hat{y}$) for Y can be given by:

- ☐ A: $\hat{y} = \operatorname{argmin}_y \sum_{i=1}^n (y - x_i)^2$
☐ B: $\hat{y} = \operatorname{argmin}_y \sum_{i=1}^n |y - x_i|$
☐ C: $\hat{y} = \frac{1}{n} \sum_{i=1}^n x_i$
☐ Both A & C
☐ Both B & C

7. (1 point) What is the result of $\lim_{x \rightarrow 0} \frac{\sin(3x)}{e^x - 1}$?

- ☐ 3 ☐ 1 ☐ 0 ☐ Undefined

8. (1 point) Given the function $f(x, y, z) = xy^2z^3$, what is $\frac{\partial^2 f}{\partial y \partial z}$?

- ☐ $2xyz^3$ ☐ $6xy^2z^2$ ☐ $6xyz^2$ ☐ $2xz^3$