Winter 2026 DSC 240: Introduction to Machine Learning
Homework 1

Due: **Monday, Jan 26th, 11:59 pm PST**

Notes:

1. **This assignment is to be done individually.** You may discuss the problems at a general level with others in the class (e.g., about the concepts underlying the question, or what lecture or reading material may be relevant), but the work you turn in must be solely your own.

2. Be aware of the late policy in the course syllabus – i.e., *no late days for all the assignments,* so **it is your responsibility to turn in your assignment to Gradescope by the due time.**

3. Justify every answer you give – **show the work** that achieves the answer or **explain** your response.

4. Any updates or corrections will be posted on Piazza, so check there occasionally.

5. It is recommended to typeset the homework in LaTeX or markdown for the submission. However, neatly, hand-written homework are also allowed.

6. This assignment requires **two parts of the submission**. Both of these submissions will be available on Gradescope, and **both submissions are MANDATORY** for receiving full credit for the assignment:

    (a) Written form of the submission as a **PDF** file format. This needs to be submitted in **Homework 1 - Report** on Gradescope.

    (b) Code submission as a **.py** python script or a **.ipynb** jupyter notebook file. This needs to be submitted in **Homework 1 - Code** on Gradescope.

7. You can complete the assignment in a Jupyter notebook, which allows you to write both the text (using LaTeX for math) and the code in one file. Once done, convert the notebook to a PDF and submit it to the **Homework 1 - Report** on Gradescope. Also, submit the .ipynb notebook file (with the code) to the **Homework 1 - Code** on Gradescope. This is just a suggestion – you can also submit the report using LaTeX/Markdown/handwriting and the code in a separate .py/.ipynb file if you prefer.

8. Be sure to re-read the **"Academic Integrity"** on the course syllabus. You must complete the section below. If you answered Yes to either of the following two questions, give corresponding full details.

Did you receive any help whatsoever from anyone in solving this assignment?
Did you give any help whatsoever to anyone in solving this assignment?

---

1. **(10 points)** Let

$$A = \begin{bmatrix} 4 & 1 & 3 & 6 \\ 2 & 7 & 5 & 3 \end{bmatrix}, B = \begin{bmatrix} 0 & 4 \\ 7 & 6 \\ 5 & 8 \\ 3 & 11 \end{bmatrix}, C = \begin{bmatrix} -13 & 0 & 2 \\ 5 & 2 & 10 \\ 0 & 7 & 9 \end{bmatrix}, D = \begin{bmatrix} 5 & -3 & -7 \\ 4 & 0 & 10 \\ 7 & 3 & 11 \end{bmatrix}, E = \begin{bmatrix} -4 & 5 \\ 12 & 7 \end{bmatrix}.$$

If possible, compute the following:

(a) $(3B)^T$

(b) $(A - B)^T$

(c) $(2B^T - A)^T$

(d) $(C + 2D^T + E)^T$

(e) $(-A)^T E$

Note: T subscript means transpose.

2. **(6 points)** Let

$$A = \begin{bmatrix} 2 & 7 & 3 \\ 1 & 0 & 9 \\ -1 & 2 & 10 \end{bmatrix}, B = \begin{bmatrix} -2 & 0 & 3 \\ 2 & -1 & 7 \\ 6 & 4 & -3 \end{bmatrix}.$$

Is $AB = BA$? Justify your answer.

3. **(8 points)** Compute the $\ell_1/\ell_2/\ell_\infty$ norms of the following vectors:

- $[0, 0, 0]$
- $[1, 2, 3]$
- $[2, 4, 6]$. How are the norms of $[2, 4, 6]$ related to those of $[1, 2, 3]$?
- Can you find a vector with negative norms, why?

4. **(6 points)** Given $X = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$ where $\mathbf{x}_i \in \mathbb{R}^m$ for all i, and $Y^T = [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n] \in \mathbb{R}^{p \times n}$ where $\mathbf{y}_i \in \mathbb{R}^p$ for all i. Show that

$$XY = \sum_{i=1}^{n} \mathbf{x}_i (\mathbf{y}_i)^T.$$

(Hint: use the definition of matrix multiplication, and change the order of summation.)

5. **(8 points)** Given $X \in \mathbb{R}^{m \times n}$, show that the matrix $X^T X$ is symmetric and positive semi-definite. When is it positive definite? (Hint: you may use the results in Q4.)

6. **(5 points)** Given $g(x, y) = e^{(x+y)} + e^{3xy} + e^{y^4}$, compute $\frac{\partial g}{\partial x}$ and $\frac{\partial g}{\partial y}$.

7. **(18 points)** Consider the matrix

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix},$$

(a) Compute the eigenvalues and corresponding eigenvectors of A. You are allowed to use PYTHON to compute the eigenvectors (but not the eigenvalues). Please include the code that you used for computing eigenvectors.

(b) What is the eigen-decomposition of A?

(c) What is the rank of A?

(d) Is A positive definite?

(e) Is A positive semi-definite?

(f) Is A singular?

8. **(12 points)** Consider the linear classifier in two dimensions: $h(\mathbf{x}) = \text{sign}(\mathbf{w}^T\mathbf{x})$ where $\mathbf{w} = [w_0, w_1, w_2]^T$ and $\mathbf{x} = [1, x_1, x_2]^T$. Technically, $\mathbf{x}$ has three coordinates, but we call this feature vector two-dimensional because the first coordinate is fixed at 1.

(a) (6 points) Show that the regions on the plane where $h(\mathbf{x}) = +1$ and $h(\mathbf{x}) = -1$ are separated by a line. If we express this line by the equation $x_2 = ax_1 + b$, what are the slope a and intercept b in terms of w_0, w_1, w_2?

(b) (6 points) Draw a picture for the cases $\mathbf{w} = [1, 2, 3]^T$ and $\mathbf{w} = -[1, 2, 3]^T$ with PYTHON code.

In more than two dimensions, the $+1$ and -1 regions are separated by a *hyperplane*, the generalization of a line.

9. **(Bonus 20 points)** In this problem, we explore the perceptron learning algorithm further with data sets of different sizes and dimensions. Please include your code (.py or .ipynb) in your submission, and include your plots and explanations in the PDF you submit.

```
1       def GenerateData(margin, number):
2       data = []
3       i = 0
4       while(i<number):
5           point = np.random.randn(2,1)
6           if point[0]+point[1]-margin>0:
7               data.append([point, 0])
8               i+=1
9           elif point[0]+point[1]+margin<0:
10              data.append([point, 1])
11              i+=1
12      return data
```

Generate the data points by the following function:

(We strongly recommend to complete this question despite it being bonus. It should be highly doable after Lecture 4. It's bonus because you will get to implement the Perceptron algorithm in MP2 already.)

(a) **(5 points)** Generate a linearly separable data set of size 20. The data space is the same as the previous question (Q8). Plot the examples $\{(\mathbf{x}_n, y_n)\}$. Find a target function f that perfectly classifies the data, and draw its decision boundary. Be sure to mark the examples from different classes differently (using different color / bullets), and add labels to the axes of the plot.

(b) **(5 points)** Run the perceptron learning algorithm on the data set above. Report the number of updates that the algorithm takes before converging. Plot the examples $\{(\mathbf{x}_n, y_n)\}$, the target function f, and the final hypothesis g in the same figure. Comment on whether f is close to g.

(c) **(3 points)** Repeat everything in (b) with another randomly generated data set of size 20. Compare your results with (b).

(d) **(3 points)** Repeat everything in (b) with another randomly generated data set of size 100. Compare your results with (b).

(e) **(4 points)** Repeat everything in (b) with another randomly generated data set of size 1,000. Compare your results with (b).

10. **(9 points)** You are asked to build a machine learning system to estimate someone's blood pressure (two numbers: systolic and diastolic; consider them to be real-valued) based on the following inputs: the patient's sex, age, weight, average grams of fat consumed per day, number of servings of red meat per week, servings of fruits and vegetables per day, smoker or non-smoker. You are given a training data set of values for all of these variables and the blood pressure numbers for 10,000 patients.

Answer (and explain) the following questions

(a) What kind of machine learning problem is this (e.g., supervised learning, unsupervised learning, reinforcement learning, ...)?

(b) What is the label space for this problem?

(c) What is the output space for this problem?

11. **(18 points)** There are 100,000 emails used to train a spam detection system – 5,000 of them are of spam and the rest are non-spam. To test the system, you have 10,000 emails – 2,000 spams and 8,000 non-spams – in your test set.

The results of the test are as follows: 250 of the spam emails are classified as non-spam, and the rest are classified as spam; 250 of the non-spam emails are classified as spam, and the rest are classified as non-spam.

(a) Show the confusion matrix for this binary classification experiment. Label it clearly and fill out the table entries.

(b) What is the false positive rate of the system in this experiment?

(c) What is the false negative rate?

(d) What is the error rate?

(e) What is the precision?

(f) What is the sensitivity?