CS 4641: Machine Learning
Homework 0 — Preview

Instructions

This is an ungraded homework assignment to help students who are uncertain if they have enough of the background material decide if they want to take this course. Much of the material in this course will build off of and expect a certain familiarity with mathematical and programming concepts from a variety of sources, and we will cover some of the background material briefly when necessary, but if these concepts are new or confusing you may have to spend additional time on top of the regular course work to fully understand the material. This is intended to be a guide, not a test, and it should not be difficult or take much time to complete. If you find yourself taking a long time to complete any of these questions, be ready to spend extra time on the related material when it is covered in class. If this is true for *several* questions, **seriously consider whether you have the time in your schedule for this class**.

Questions

1. Programming.
   (a) Install Python, NumPy, Matplotlib into a virtualenv using `pip`. (*you will use these libraries for the rest of this course*)
   (b) Write a python class which has one method which returns the number of times the method has been called since the object was instantiated.
   (c) Write a function which takes as arguments: another function, a list of arguments, and returns the output of the given function and the time it took for the passed in function to run given those arguments.
   (d) Write a function that takes a dictionary (`dict1`) and returns a dictionary containing the number of keys in `dict1` with the key `'num_keys'`, a list of the unique data types of the values in `dict1` with the key `'datatypes'`, and `dict1` with the key `'data'`.
   (e) Write a one line expression that builds a dictionary from a list of 2-element tuples such that the first element of each tuple is a key and the second element is the corresponding value (don't worry about duplicates).
   (f) Write a one line expression that builds a list of 2-element tuples from a dictionary *without* using `items(...)`.
   (g) Write a slice expression to get a single row or column from a numpy array.
   (h) Write a slice expression to get all the rows of a numpy array where one of the values is larger than the mean of its column plus 2 times the standard deviation of its column.
2. Linear Algebra

(a) What is the rank of this matrix:

$$\begin{bmatrix} 1 & 2 & -3 \\ -1 & -3 & 5 \\ 3 & 0 & 3 \end{bmatrix}$$

(b) What are the eigenvalues and eigenvectors of this matrix:

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix}$$

(c) Prove that if $\mathbf{A}$ and $\mathbf{B}$ are positive definite matrices of the same size, so is the matrix $\mathbf{C} = \mathbf{A} + \mathbf{B}$.

3. Calculus

(a) What is the maximum value of the following function, and at what x value is it obtained?

$$f(x) = -x^4 + \frac{4}{3}x^3 + 4x^2 + 3$$

(b) What is the partial derivative of the following function with respect to y?

$$f(x, y) = 4x^2y^2 - \frac{1}{2}x^2y + \frac{2}{3}y^2x - 2xy + 4$$

(c) What does the following expression evaluate to?

$$\int_0^{2\pi} \left(\int_0^{\infty} e^{-\frac{r}{2}^2} r dr \right) d\theta$$

4. Probability and Statistics

(a) If X is a random variable that takes values in the interval $[0, \infty)$ with probability density $p(X) = e^{-X}$, and $f(X) = 4 - X^2$, what is the expected value of $f(X)$?

(b) For the same random variable, what is the probability that X is 0.5? ($P(X = 0.5)$)

(c) For the same random variable, what is the probability that X is between 0.25 and 0.75? ($P(0.25 \leq X \leq 0.75)$)

(d) If Y is a second random variable, independent of X, in the interval $[0, 10]$ with uniform density, what is the expected value of $Y \cdot f(X)$?

(e) Say you've invented a brand new ML algorithm and you want to test its performance emperically. You run a bunch of experiments (say, N of them) measuring the performance, P, and each time get a slightly different result. Treating P as a random variable, what should you say is the expected performance of your algorithm? What are the 95% confidence intervals for this value? What assumptions did you make to get these values? Explain what these confidence intervals represent.