Morales, Samantha
CSCI 3203: Tashfeen's Machine Learning I

Homework 1

Please show all work for full credit. You're only allowed to use the standard built-in Python 3 libraries.

Question 1. Watch Mary's Room: A philosophical thought experiment – Eleanor Nelsen on YouTube.
Do you think there is something about human learning that is different than machine learning?

Not intrinsically, I believe that human learning relies heavily on perception and up until now I don't think we've really implemented machine learning algorithms that can truly perceive and learn in the same way humans can; however, I don't think this would be impossible to do.

Question 2. Rate your skills out of ten in the following subjects,

1 Probability and Statistics - 5
2 Calculus - 5
3 Linear Algebra - 5
4 Python - 7

Question 3. Of the sub-plots (a), (b), (c) and (d) in figure ??, state which ones have no correlation, negative correlation, positive correlation and non-linear correlation.

C, B, A, D; Respectively

Question 4. Assume that the probability of Alice going to class everyday given she got a good grade is $\frac{5}{6}$ and the probability of her getting a good grade is $\frac{2}{5}$ while the probability of her going to class everyday is $\frac{1}{3}$. What is the probability that Alice gets a good grade given she went to class everyday?
Hint: Thomas Bayes.

$$P(B|A) = \frac{5}{6}, P(A) = \frac{2}{5}, P(B) = \frac{1}{3}$$
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{\frac{5}{6} \times \frac{2}{5}}{\frac{1}{3}} = \frac{\frac{1}{3}}{\frac{1}{3}} = 1$$

Question 5. We define the natural numbers as,

$$\mathbb{N} = \{1, 2, 3, 4, 5, ...\}$$

Let x be the arithmetic mean of the first $2^{33} - 1$ natural numbers. What is $\log_2(x)$?
Hint: Carl Friedrich Gauss.

$$\frac{\frac{n(n+1)}{2}}{n} = \frac{n(n+1)}{2n} = \frac{n+1}{2} = \frac{2^{33} - 1 + 1}{2} = 2^{32} => \log_2(2^{32}) = 32$$

Question 6. Let `arr = [x % 2 for x in range((2**100)-1)]` be a Python list using Python list comprehension. What is the statistical mode of the list `arr`? Justify your answer.

0

Question 7. Define a function $f : \mathbb{R} \to \mathbb{R}$,

$$f(x) = \frac{\sin(x^2)}{y}$$

Calculate the following,

- The partial derivative with respect to y, $\frac{\partial f}{\partial y}$

$$-\frac{\sin(x^2)}{y^2}$$

- The partial derivative with respect to x, $\dfrac{\partial f}{\partial x}$

$$\frac{2x\cos(x^2)}{y}$$

- The gradient vector $\nabla f(x, y)$.

$$[\frac{2x\cos(x^2)}{y}, -\frac{\sin(x^2)}{y^2}]$$

Question 8. Let $\vec{v_1} = [e, \pi, \sqrt{2}]$ and $\vec{v_2} = [1, 2, 0]$. What is the dot product $\vec{v_1} \cdot \vec{v_2}$? Please do not give an approximation.

$e + 2\pi$

Question 9. We define a matrix of n rows and p columns with real values as $\mathbf{A} \in \mathbb{R}^{n\times p}$. We can think of $\mathbf{A}$ as having n row vectors with the $i^{\text{th}} \leq n$ row vector being $\text{row}(\mathbf{A})_i$. Similarly, we can think of $\mathbf{A}$ as having p column vectors with the $j^{\text{th}} \leq p$ column vector being $\text{col}(\mathbf{X})_j$.

Then matrix multiplication between $\mathbf{A} \in \mathbb{R}^{n\times p}$ and $\mathbf{B} \in \mathbb{R}^{p\times m}$ is defined as,

$$\mathbf{AB} = \begin{bmatrix} \text{row}(\mathbf{A})_1 \,\text{col}(\mathbf{X})_1 & \text{row}(\mathbf{A})_1 \,\text{col}(\mathbf{X})_2 & \cdots & \text{row}(\mathbf{A})_1 \,\text{col}(\mathbf{X})_m \\ \text{row}(\mathbf{A})_2 \,\text{col}(\mathbf{X})_1 & \text{row}(\mathbf{A})_2 \,\text{col}(\mathbf{X})_2 & \cdots & \text{row}(\mathbf{A})_2 \,\text{col}(\mathbf{X})_m \\ \vdots & \vdots & \ddots & \vdots \\ \text{row}(\mathbf{A})_n \,\text{col}(\mathbf{X})_1 & \text{row}(\mathbf{A})_n \,\text{col}(\mathbf{X})_2 & \cdots & \text{row}(\mathbf{A})_n \,\text{col}(\mathbf{X})_m \end{bmatrix}$$

where $\text{row}(\mathbf{A})_i \,\text{col}(\mathbf{X})_j \in \mathbb{R}$ is a dot or inner product.

Further, the transpose of $\mathbf{A} \in \mathbb{R}^{n\times p}$ is written as $\mathbf{A}^T \in \mathbb{R}^{p\times n}$ and is defined as,

$$\text{If } \mathbf{A}^T \text{ is the transpose of } \mathbf{A} \text{ then } \text{row}(\mathbf{A})_i = \text{col}(\mathbf{A}^T)_j.$$

(a) What must be true of the number of columns of $\mathbf{A}$ and number of rows of $\mathbf{B}$ for $\mathbf{AB}$ to be defined?

The number of columns in A must match the number of rows in B.

(b) What information do the number of rows of $\mathbf{A}$ and number of columns of $\mathbf{B}$ give you about the dimensions of the product $\mathbf{AB}$?

The number of rows in A and the number of columns in B give us the number of rows and columns in AB (respectively).

(c) For the two matrices bellow, give the product $\mathbf{AB}$.

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 4 & 2 & 2 \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} 5 & 4 & 3 \\ 8 & 9 & 5 \\ 6 & 5 & 3 \\ 11 & 9 & 6 \end{bmatrix}$$

(d) Prove that the matrix product is not a symmetric relation, i. e., $\mathbf{AB} \neq \mathbf{BA}$. Hint: can you construct a small counter example? BA is not even possible because the number of columns in B does not match the number of rows in B.

(e) Prove that $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

Proof. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$

(1) $(\mathbf{AB})^T_{j,i} = \mathbf{AB}_{i,j} = \text{row}(\mathbf{A})_i \,\text{col}(\mathbf{B})_j$

(2) $(\mathbf{B}^T\mathbf{A}^T)_{j,i} = \text{row}(\mathbf{B}^T)_j \,\text{col}(\mathbf{A}^T)_i = \text{row}(\mathbf{A})_i \,\text{col}(\mathbf{B})_j$

(3) therefore, $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$

QUOD ERAT DEM■

(f) For $\mathbf{X} \in \mathbb{R}^{n\times p}, \beta \in \mathbb{R}^p, y \in \mathbb{R}^n$ prove that,

$$\beta^T \mathbf{X}^T y = y^T \mathbf{X} \beta$$

Proof. $\beta^T \mathbf{X}^T y = y^T \mathbf{X} \beta$

(1) $\beta^T \mathbf{X}^T y$ and $y^T \mathbf{X} \beta$ return a 1x1 matrix
(2) we can perform a transpose on either of them since and 1x1 matrix equals itself when transposed
(3) $(\beta^T \mathbf{X}^T y)^T = y^T (\beta^T \mathbf{X}^T)^T$
(4) $y^T (\beta^T \mathbf{X}^T)^T = y^T \mathbf{X} \beta$
(5) therefore, $\beta^T \mathbf{X}^T y = y^T \mathbf{X} \beta$

QUOD ERAT DEM■

(g) Let $f(\beta) = \beta^T \beta$. Prove that the derivative $\frac{\mathrm{d}f}{\mathrm{d}\beta} = 2\beta$.

Proof. $\frac{\mathrm{d}f}{\mathrm{d}\beta} = 2\beta$, for $f(\beta) = \beta^T \beta$

(1) Assuming the definition for β as defined in part f, $\beta \in \mathbb{R}^p$
(2) $\mathrm{col}(\beta)_i = \mathrm{row}(\beta^T)_i$
(3) since the values of β are being muliplied with themselves, this means $\beta^T \beta = \langle \beta_1^2 + \beta_2^2 ... + \beta_i^2 \rangle$
(4) since we have to take the partial dervative for each β_i every time, we are left with only the β_i since the rest are considered constants
(5) this leaves us with 2β
(6) therefore, $\frac{\mathrm{d}f}{\mathrm{d}\beta} = 2\beta$, for $f(\beta) = \beta^T \beta$

QUOD ERAT DEM■

Question 10. For $x \in \mathbb{R}$ we have,

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

Give,

$$\min_x f(x) = \min_x \left(\frac{x^4}{4} - \frac{x^3}{3} \right)$$

Plot f to double check your answer.

the global minimum of $f(x) = -\frac{1}{12}$ occurs at $x = 1$

Question 11. Let $x \sim \mathcal{N}(0, 1)$ be a normally distributed random variable with mean 0 and standard deviation 1. What is the likelihood that $x = \sqrt{2}$. Give an exact answer.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

$$f(x) = \frac{1}{1\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-0}{1})^2} = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

$$f(\sqrt{2}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-1} = \frac{1}{e\sqrt{2\pi}}$$

or 0 idk. I don't think x can be any specific value so it's 0 I think, but also the answer from before is the height of the graph at $f(\sqrt{2})$

Question 12. The code snippet in listing 1 reads the plain text of the novel *The Cosmic Computer* by the famous American science fiction writer Henry Beam Piper over the internet and saves it to a string variable `text`.

```python
from urllib.request import urlopen as get

url = 'https://www.gutenberg.org/files/20727/20727.txt'
with get(url) as response:
```

```
    text = response.read().decode('utf-8')

print(text)
```

LISTING 1. A Python program to download the text of the *The Cosmic Computer* by Piper.

Write a Python program that prints out the ten most used words in the novel that have more than 5 letters. State your findings. Put your code in a file called `frequency.py`.

maxwell: 86
project: 76
anything: 74
something: 72
storisende: 65
poictesme: 61
everybody: 61
jacquemont: 59
command: 58
started: 57

Question 13. Watch Laziness in Python - Computerphile on YouTube.

Implement the function `fibonacci(n)` in the code listing 2.

```
def fibonacci(n):
    # implement me

for t in fibonacci(50):
    print(t)
```

LISTING 2. A Python 3 program to print out the first $n \geq 0$ Fibonacci numbers.

Run your program for $n = 50$. What are the last five numbers printed? Put your code in a file called `fibonacci.py`.

1134903170, 1836311903, 2971215073, 4807526976, 7778742049

SUBMISSION INSTRUCTIONS

1 Submit a PDF that answers all the questions.
2 Submit Python files, e. g., `frequency.py` and `fibonacci.py`.

OKLAHOMA CITY UNIVERSITY, PETREE COLLEGE OF ARTS & SCIENCES, COMPUTER SCIENCE