CSE514, Fall 2021, HW 1 Name: Anh Le Student ID: 488493

Note: This homework is worth a total of 15 points

Q1: Give two examples of the following data types:

Q1a: (2pts) Numerical, Discrete:

- Number of students in a class, number of cars in a parking lot

Q1b: (2pts) Numerical, Continuous:

- Temperature in a day, weight of a person

Q1c: (2pts) Categorical, Nominal:

- Gender (M/F), race

Q1d: (2pts) Categorical, Ordinal:

- Military rank, flight classes

Q2: (2pts) Prove that the function

f(*x*): count the number of zero values in x

is not a mathematical norm by giving a counter-example to two of the three properties of a norm:

Q2a: $f(x) > 0$ for $x \in V$ and $x \neq 0$

Q2b: $f(x + y) \leq f(x) + f(y)$ for $x, y \in V$

$$\left.\begin{array}{l} x = [1, -1] \\ y = [-1, 1] \end{array}\right\} \rightarrow f(x+y) = f([0,0]) = 2 > f(x) + f(y) = 0 + 0 = 0$$

Q2c: $f(\lambda x) = |\lambda| f(x)$ for all $\lambda \in$ R and $x \in V$

$$\left.\begin{array}{l} x = [1] \\ \lambda = 0 \end{array}\right\} \rightarrow f(\lambda x) = f([0]) = 1 \neq |\lambda| f(x) = 0$$

Q3: (1pt) Show that a vector's dot product with itself is also the square of it's Euclidean norm

- Euclidean norm:

$$\|x\|_2 = \left(\sum_{i=1}^{K} |x_i|^2\right)^{1/2} = \sqrt{x_1^2 + x_2^2 + \dots + x_K^2} \quad (1)$$

- Dot product of vector $x = [x_1, x_2, \dots, x_K]$

$$x^T x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix} [x_1 \ x_2 \ \dots \ x_K] = x_1^2 + x_2^2 + \dots + x_K^2 = \left(\sqrt{x_1^2 + x_2^2 + \dots + x_K^2}\right)^2 \quad (2)$$

$$(1)(2) \Rightarrow x^T x = \|x\|_2^2$$

Q4: A friend informs you that a casino is using loaded dice, such that:

$$p_! = prob(roll = k) = \begin{matrix} 1/12 & if\ k \in \{1,2,3\} \\ 1/4 & if\ k \in \{4,5,6\} \\ 0 & otherwise \end{matrix}$$

Q4a: (2pt) What is the entropy of a roll at this casino? **Show your work**

$$H(roll) = \sum p_i \log_2\left(\frac{1}{p_i}\right) = 3\times\left(\frac{1}{12}\log_2(12)\right) + 3\times\left(\frac{1}{4}\log_2(4)\right) \approx 2.396$$

Q4b: (2pts) How much information did you gain compared to the assumption that the dice are fair? (i.e. What's the KL divergence of the loaded dice distribution from the distribution of fair dice?) **Show your work**

Fair dice: $P_{1,2,3,4,5,6} = \frac{1}{6}$

Loaded dice: $P_{1,2,3} = \frac{1}{12}$, $P_{4,5,6} = \frac{1}{4}$

↳ $D_{KL}(\text{loaded} \,\|\, \text{fair}) = \sum_K P_{loaded} \log \frac{P_{loaded}}{P_{fair}}$

$$= 3\times\left(\frac{1}{12}\log\left(\frac{1/12}{1/6}\right)\right) + 3\times\left(\frac{1}{4}\log\left(\frac{1/4}{1/6}\right)\right) = 0.0568$$