CS561 Activity 2

Punnawish Thuwajit (KK)

September 12, 2023

1. Classification

We're given

$$f(x) = \begin{cases} x \geq 10 & A \\ x \in [8, 10) & B \\ x \in [6, 8) & C \\ x \in [4, 6) & D \\ x \in [0, 4) & E \end{cases}$$

1.1 Problem (a)

We would predict $f(5) = D$

1.2 Problem (b)

The possible outcomes are $\mathcal{Y} = \{A, B, C, D, E\}$

1.3 Problem (c)

The input features are $\mathcal{X} = \mathbb{R}_0^+$ which denotes the number of hours.

1.4 Problem (d)

The new input features are $\mathcal{X}' = \mathbb{R}_0^+ \times \{0, 1\}$ which denotes the tuples containing the number of hours and the binary indicator on whether or not the student achieves a perfect score on HW1.

1.5 Problem (e)

We define the new function $f' : \mathcal{X}' \to \mathcal{Y}$ as such: $f'(x, c) = f(x + 3.5c)$

2. Regression

2.1 Problem (a)

The possible outcomes are $\mathcal{Y} = \{a + b\cos\left(\frac{2\pi x}{T} + \theta\right) : x \in \{1, 2, \ldots, 365\}\}$

2.2 Problem (b)

We're basically given pairs of (x, y). We can use curve fitting algorithms (least square) to fit for the best values of a, b, T, θ.

2.3 Problem (c)

Two points (we would end up with a linear system of equations with 2 variables), under the assumption the function correctly describes our desired relation.

2.4 Problem (d)

365, since all 365 points are independent.

2.5 Problem (e)

The second model requires all 365 points. In contrast, the first would require far less. However, the fist model forces an assumption that the function follows a cosine curve (which may not be the case).