Northeastern University
College of Engineering
Group member name(s): Yanming Liu
Group member UID(s): 002199402

May 27, 2022
Assignment 6

IE 7374: Machine Learning

Prove if this is a convex function

Ex: $P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$

$$\left.\begin{aligned} &D = (x_1, x_2, \dots, x_n) \\ &P(X_i = x_i) = \theta^{x_i}(1-\theta)^{1-x_i} \end{aligned}\right\} \Rightarrow P_\theta(D) = \theta^{\sum x_i}(1-\theta)^{N-\sum x_i}$$

we want to calculate $\hat{\theta}_{ML}$

$$\hat{\theta}_{ML} = \arg\max_\theta P_\theta(D) = \arg\max_\theta \theta^{\sum x_i}(1-\theta)^{N-\sum x_i} \xRightarrow[\log]{\text{Apply}} \arg\max_\theta \log P_\theta(D) = \arg\max_\theta \log\left(\theta^{\sum x_i}(1-\theta)^{N-\sum x_i}\right)$$

$$= \sum x_i \log\theta + (N - \sum x_i)\log(1-\theta) \longrightarrow \text{is it convex?}$$

Answer to Question

Let $f(\theta) = \sum x_i log\theta + (N - \sum x_i)log(1-\theta)$,
Derive the function on θ, $f'(\theta) = \frac{\sum x_i}{\theta\ln(2)} - \frac{N}{(1-\theta)\ln(2)} + \frac{\sum x_i}{(1-\theta)\ln(2)}$
Derive twice,

$$\begin{aligned} f''(\theta) &= \frac{-\sum x_i}{\theta^2\ln(2)} - \frac{N}{(1-\theta)^2\ln(2)} + \frac{\sum x_i}{(1-\theta)^2\ln(2)} \\ &= \frac{-(1-\theta)^2\sum x_i - N\theta^2 + \theta^2\sum x_i}{\theta^2(1-\theta)^2\ln(2)} \\ &= \frac{2(\theta-1)\sum x_i - N\theta^2}{\theta^2(1-\theta)^2\ln(2)} \\ &< \frac{-N(2\theta-1) + N\theta^2}{\theta^2(1-\theta)^2\ln(2)} \\ &= \frac{-N(\theta-1)^2}{\theta^2(1-\theta)^2\ln(2)} < 0, when\ 0 < \theta < 1 \end{aligned}$$

So, it is not a convex function, while it is a concave one.