Homework week 5

Nguyễn Minh Đức
Student code: 11204838

Octorber 2022

1 Question 1

Let X = $\begin{bmatrix} 1 & x_{11} & x_{12} & ... & x_{1d} \\ 1 & x_{21} & x_{22} & ... & x_{2d} \\ ... & ... & ... & ... & ... \\ 1 & x_{n1} & x_{n2} & ... & x_{nd} \end{bmatrix} = \begin{bmatrix} x_1^T \\ x_2^T \\ ... \\ x_n^T \end{bmatrix}$; W = $\begin{bmatrix} w_0 \\ w_1 \\ ... \\ w_d \end{bmatrix}$;

Y = $\begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_n \end{bmatrix}$; Ŷ = $\begin{bmatrix} \hat{y_1} \\ \hat{y_2} \\ ... \\ \hat{y_n} \end{bmatrix} = \begin{bmatrix} h_\theta(x_1^T W) \\ h_\theta(x_2^T W) \\ ... \\ h_\theta(x_n^T W) \end{bmatrix} = \theta(XW)$

We have:

$$L = -(y^T log\hat{y} + (1-y)^T log(1-\hat{y}))$$

$$\frac{\partial L}{\partial W} = (\frac{\partial \hat{y}}{\partial W})^T \frac{\partial L}{\partial \hat{y}}$$

$$\frac{\partial \hat{y}}{\partial W} = \frac{\partial \theta(XW)}{\partial W} = \frac{\partial \theta(XW)}{\partial XW}\frac{XW}{\partial W} = \frac{\partial \theta(XW)}{\partial XW} X$$

$$h_\theta(x) = \frac{1}{1+e^{-x}}$$

$$\frac{\partial h_\theta(x)}{\partial x} = \frac{e^{-x}}{(1+e^{-x})^2} = \frac{1}{1+e^{-x}} * \frac{e^{-x}}{1+e^{-x}} = (1-h_\theta(x))h_\theta(x)$$

$$\frac{\partial \theta(XW)}{\partial W} = [[1-\theta(XW)] \odot \theta(XW)]^T X = (1-\hat{y}) \odot \hat{y} X$$

$$\frac{\partial L}{\partial \hat{y}} = -(y^T \frac{1}{\hat{y}} + (1-y)^T \frac{-1}{1-\hat{y}})$$

$$
\begin{aligned}
\frac{\partial L}{\partial W} &= [(1-\hat{y}) \odot \hat{y}X]^T((1-y)^T\frac{1}{1-\hat{y}} - y^T\frac{1}{\hat{y}}) \\
&= X^T(1-\hat{y}) \odot \hat{y} * ((1-y)^T\frac{1}{1-\hat{y}} - y^T\frac{1}{\hat{y}}) \\
&= X^T(\hat{Y}^T(1-Y) - (1-\hat{Y})^TY) \\
&= X^T(\hat{Y} - \hat{Y}^TY - Y + \hat{Y}^TY) \\
&= X^T(\hat{Y} - Y)
\end{aligned}
$$

2 Question 2

Logistic model for loan approved prediction:

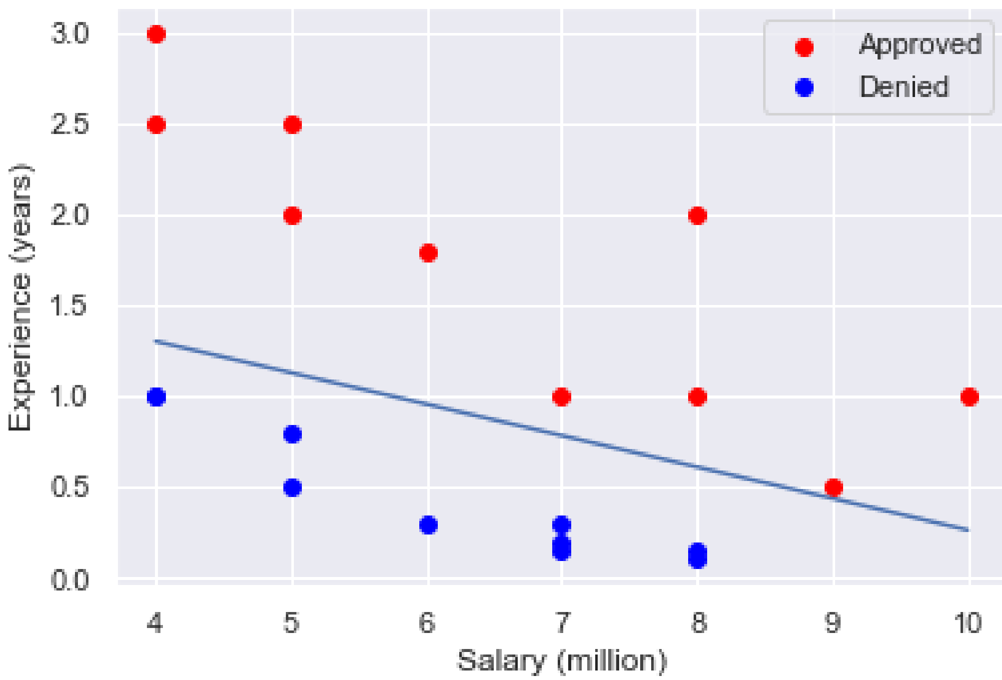
$$
\theta = \frac{1}{1 + e^{-(-7.252 + 0.629 * \text{Lương} + 3.642 * \text{Thời gian làm việc})}}
$$

3 Question 3

Assume t is decision boundary

$$
\begin{aligned}
\Rightarrow \hat{y} &\geq t \\
\Rightarrow \frac{1}{1+e^{-X^TW}} &\geq t \\
\Rightarrow \frac{1}{t} &\geq 1 + e^{-X^TW} \\
\Rightarrow \frac{1}{t} - 1 &\geq e^{-X^TW} \\
\Rightarrow ln(\frac{1}{t} - 1) &\geq -X^TW \\
\Rightarrow X^TW &\geq -ln(\frac{1}{t} - 1)
\end{aligned}
$$

Decision boundary for Question 2



4 Question 4

Binary Cross Entropy loss:

$$H = \frac{\partial^2 L}{\partial^2 W} = X^T \hat{y}(1-\hat{y})X$$

We have: $\hat{y} \geq 0$ and $(1-\hat{y}) \geq 0 \Rightarrow \hat{y}(1-\hat{y}) \geq 0 \; \forall v \in R$

$$\begin{aligned} v^T H v &= v^T X^T \hat{y}(1-\hat{y}) X v \\ &= v^T X^T ((\hat{y}(1-\hat{y}))^{\frac{1}{2}})^2 X v \\ &= ||(\hat{y}(1-\hat{y}))^{\frac{1}{2}} X v||_2^2 \geq 0 \forall \text{v} \end{aligned}$$

$\Rightarrow$Positive semi-definite $\Rightarrow$ loss function in convex
MSE loss:

$$L = ||y - \hat{y}||_2^2 \text{ where } \hat{y} = \frac{1}{1+e^{-X^T\theta}}$$

$$\begin{aligned} \frac{\partial L}{\partial W} &= -2(y-\hat{y})\hat{y}(1-\hat{y})X \\ &= -2(y\hat{y} - y\hat{y}^2 - \hat{y}^2 + \hat{y}^3)X \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 L}{\partial^2 W} &= -2(y - 2y\hat{y} - 2\hat{y} + 3\hat{y}^2)XX\hat{y}(1-\hat{y}) \\
&= -2[3\hat{y}^2 - 2\hat{y}(y+1) + y]X^2\hat{y}(1-\hat{y}) \\
\text{If y = 0} & \\
H &= -2[3\hat{y}(\hat{y} - \frac{2}{3})] \geq 0 \forall \hat{y} \in [\frac{2}{3}, 1] \\
\text{If y = 1} & \\
H &= 2[3(\hat{y} - \frac{1}{3})(\hat{y} - 1)] \geq 0 \forall \hat{y} \in [\frac{1}{3}, 1]
\end{aligned}$$

$\Rightarrow$ H is not semi-definite $\Rightarrow$ MSE loss is not convex