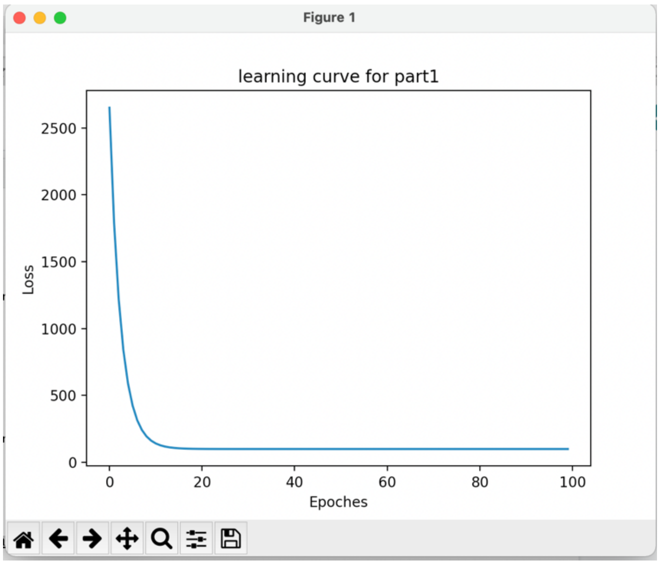
NYCU Introduction to Machine Learning, Homework 1

109550116 楊傑宇

Part. 1, Coding (60%):

Linear regression model

1.

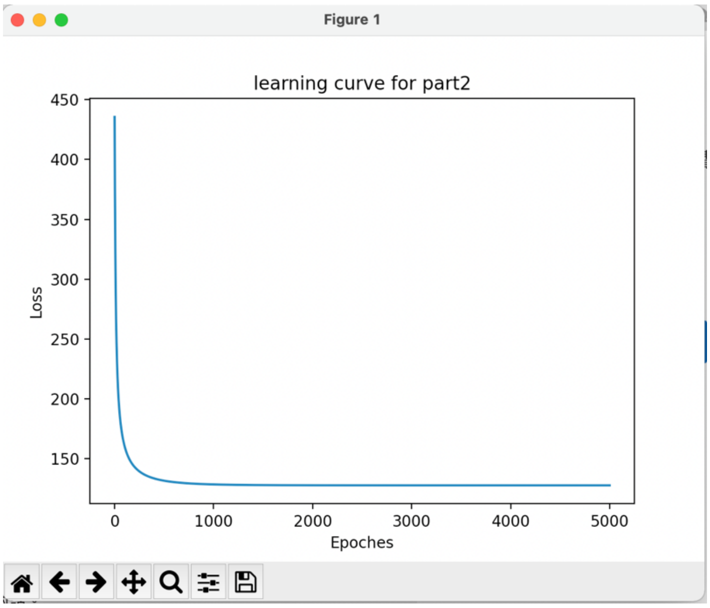


2. Error of test: 110.43819236554506
3. Weight: 52.743540399099864

   Intercepts: -0.3337589091638858

Logistic regression model

1.



2. Error of test: 47.234058602593905
3. Weight: 4.872961592425304

   Intercepts: 1.7092277751255238

Part. 2, Questions (40%):

1.

- Gradient Descent: 一次 train 全部的資料，壞處是會跑很久，好處是每次的 gradient 會很準確。
- Stochastic:一次 train 一筆資料，好處是他會跑的很快，壞處是每次的 gradient 不一定是準確的。
- Mini-batch:一次 train m 筆資料，這個方法不但跑得快，gradient 也不會與正確的差太多。

2.

- High-learning rate:高學習率因為每次變化的量很大，所以會使 loss 很快就 converge，但也可能因為變化的量太大而跳出凹槽，使 loss 變大。
- Low-learning rate:低學習率因為每次變化的量太少，所以每次學習只進步一點點，要經過很多輪才能 converge
- 因此要找到適合的 learning rate 才能使 train 又快又準確。

3.

3-1

$$\sigma(a) = \frac{1}{1+\exp(-a)}$$

$\sigma(-a)$

左式 : $\frac{1}{1+\exp(a)}$

右式 : $1-\sigma(a) = 1-\frac{1}{1+\exp(-a)}$

$$= \frac{\exp(-a)}{1+\exp(-a)} = \frac{1}{\frac{1}{\exp(-a)}+1}$$

$$= \frac{1}{\exp(a)+1}$$

因此左式 = 右式 得證

3-2

$$\sigma(x) = y = \frac{1}{1+\exp(-x)}$$

$$\Rightarrow y(1+\exp(-x)) = 1$$

$$\Rightarrow y + y\exp(-x) = 1$$

$$\Rightarrow y\exp(-x) = 1-y$$

$$\Rightarrow \exp(-x) = \frac{1-y}{y}$$

取ln

$$\Rightarrow -x = \ln\left(\frac{1-y}{y}\right)$$

$$\Rightarrow x = \ln\left(\frac{y}{1-y}\right)$$

$$\Rightarrow \sigma^{-1}(x) = \ln\left(\frac{x}{1-x}\right).$$ 得證

4.

4

$$\frac{\partial E}{\partial w} = \frac{\partial E}{\partial y}\frac{\partial y}{\partial a}\frac{\partial a}{\partial w} = \sum\left(\frac{\partial(-t\log y-(1-t)\log(1-y))}{\partial y} \times \frac{\partial y}{\partial a} \times \frac{\partial a}{\partial w}\right)$$

$$= \sum\left[\left(\left(\frac{\partial(-t\log y)}{\partial y}\right) + \left(\frac{\partial(-(1-t)\log(1-y))}{\partial y}\right)\right) \times y(1-y) \times \phi\right]$$

$$= \sum\left[\left(\frac{-t}{y} + \frac{1-t}{1-y}\right) \times y(1-y) \times \phi\right]$$

$$= \sum\left(\frac{y-t}{y(1-y)} \times y(1-y) \times \phi\right) = \sum(y-t)\phi$$

得證 #