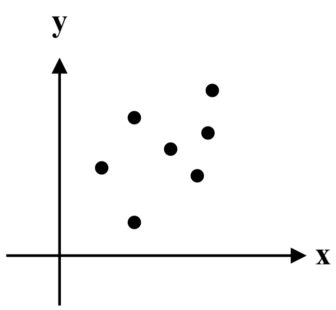
SC201 Lecture 4

Supervised Learning

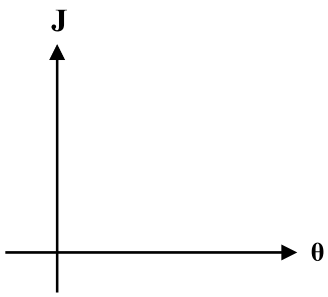
< ______________ regression >	**< ______________ regression >**

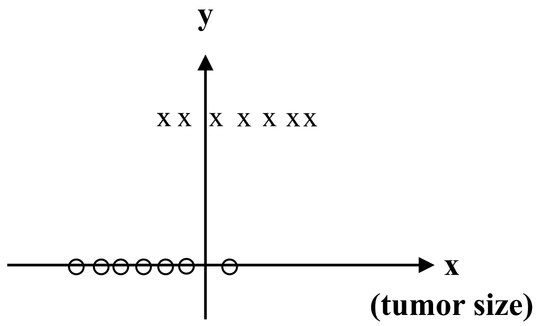


hi = ___________

J = ________________

θ = __________________________



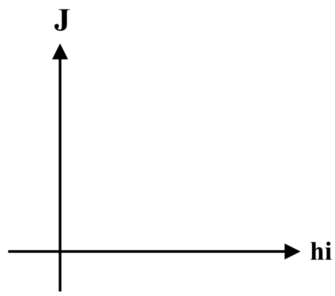


hi = ___________

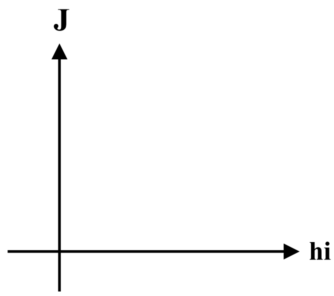
J = ______________________________

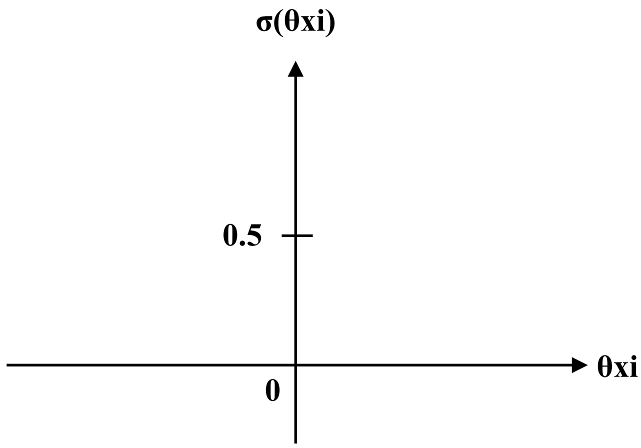
θ = __________________________

【yi==0】



【yi==1】

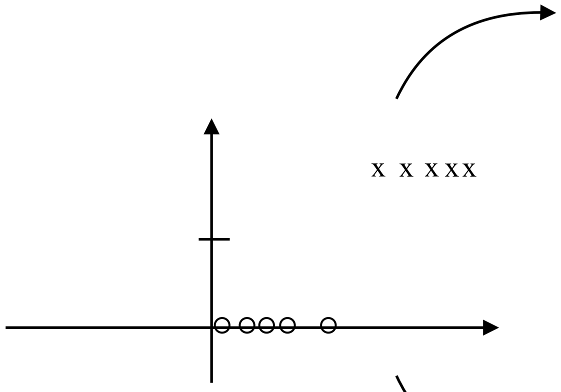




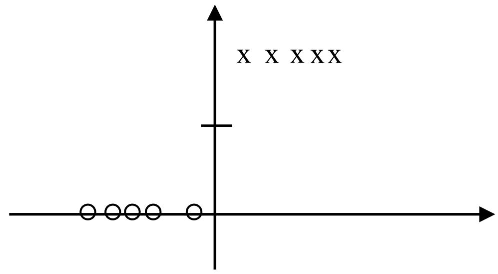
θxi _________ , σ(θxi) ___________

θxi _________ , σ(θxi) ___________

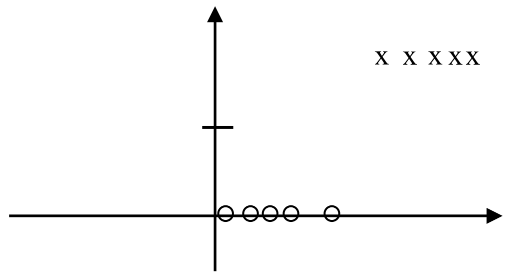
< Tumor size problem >



① ______________________________



② ______________________________



---

visualize.py

- **It is possible to raise xi to exponent (______________________________)**

feature xi → θxi+b → ______________ → __________________

feature xi, xi^2 → _______________ → ___________________ → __________________

< Maximum Likelihood Estimation >

The origin of ______________________________

_____ 是預測 yi == ____ 的機率

__________ 是預測 yi == ____ 的機率

➡ ____________________________

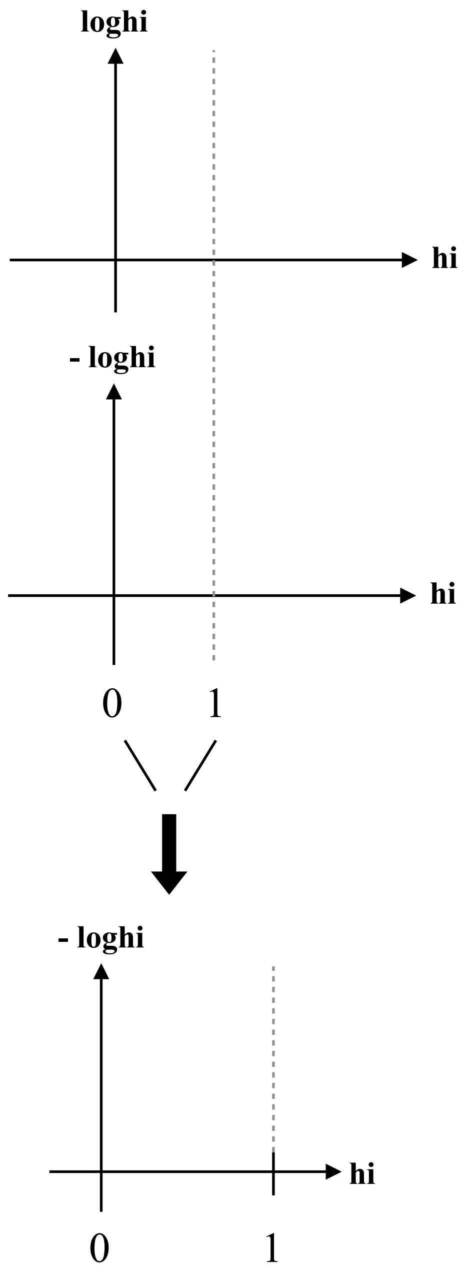
- 機率最大化

Maximize $(hi)^{yi}(1\text{-}hi)^{1\text{-}yi}$ = Maximize _____________________

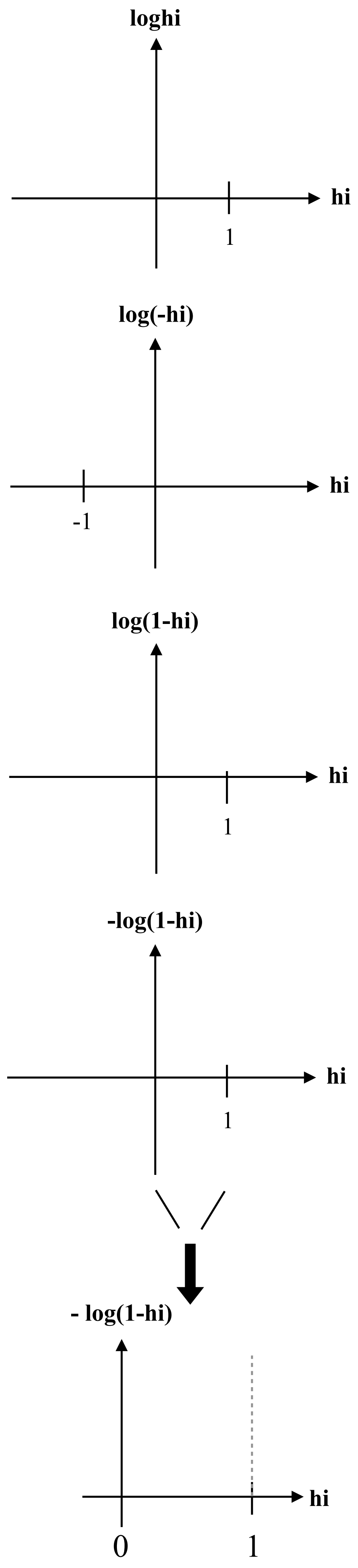
= Maximize _________________________ = Maximize _________________________

= Minimize __________________________ ➡ **J** = ____________________________

< L = -loghi when yi==1 >



< L = -log(1-hi) when yi==0 >



Gradient Descent

< Batch Gradient Descent > BGD

$$\theta = \theta - \alpha\left(\frac{d}{d\theta}\right)$$

where J = ______________________________

$$\frac{dj}{d\theta} = \underline{\hspace{6cm}}$$

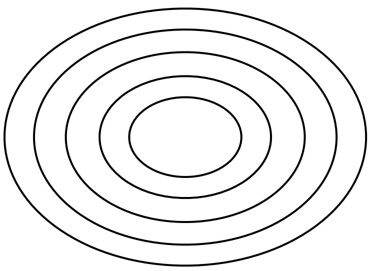
- Compute the gradient using

  ____

- Move ______________ towards the optimum
- ____________ update

  (________________________________)

- ____________ RAM needed
- ________________________



< Stochastic Gradient Descent > SGD

$$\theta = \theta - \alpha\left(\frac{d}{d\theta}\right)$$

where L = ______________________________

$$\frac{dL}{d\theta} = \underline{\hspace{6cm}}$$

- Compute the gradient using

  ____

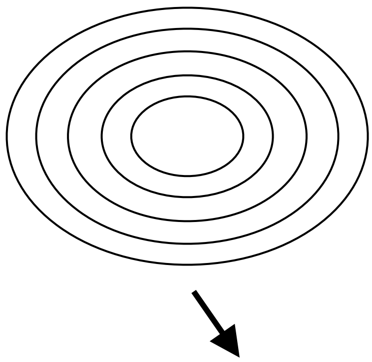
- ____________ learning process

  (often __________ from optimum)

- ____________ update

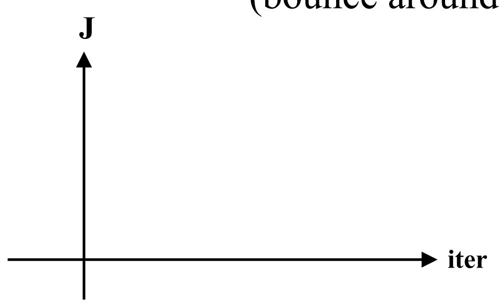
  (__________________________________)

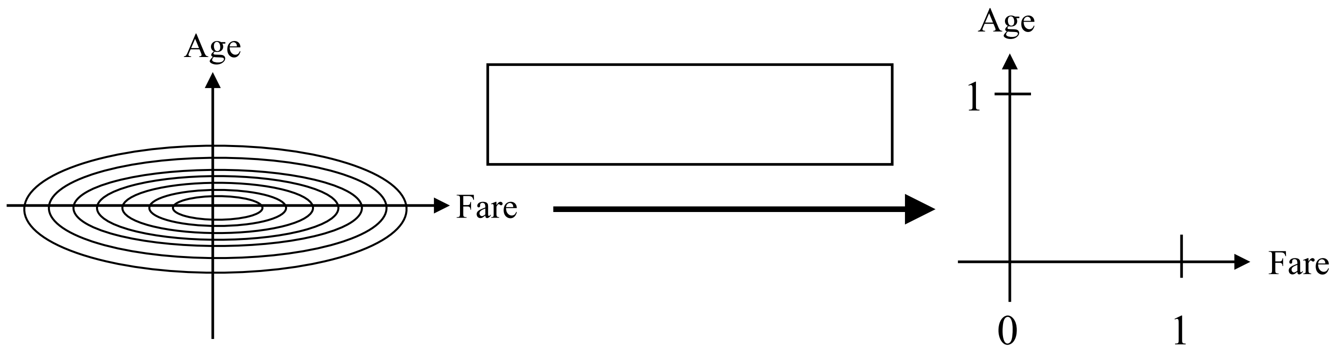
- ____________ RAM
- __________________



____

(bounce around min)





Normalization

$$X = \underline{\qquad\qquad\qquad\qquad}$$

Define Model (hi) hi = ____

↓

Define Loss Function L = ______________________________

↓

Find the best parameters (weights)

W1 = ______________________

W2 = ______________________

W3 = ______________________

$$\frac{dL}{dWi} = \underline{\qquad\qquad\qquad\qquad}$$