

ML-Lecture1-Regression&

ML-Lecture-1 Regression

Step 1: Model

$$y = b + w \cdot x_{cp}$$

\downarrow \uparrow
 parameters input

linear model: $y = b + \sum w_i x_{wi}$

\downarrow \downarrow
 bias input feature weight

Step 2: Goodness Function. (Pick the "Best" Function)

→ 搜集 Data.

loss function L : input: a function
output: how bad it is

$$L(f) = L(w, b) = \sum (y - (b + w \cdot x_{cp}))^2$$

Gradient Descent:

$\frac{dL}{dw}$ Negative \Rightarrow Increase w
Positive \Rightarrow Decrease w

→ 增加/減少決定於學習率 (learning rate)

Model Selection.

→ 避免 overfitting. (more complex model does not always lead to better performance on testing data)

Regularization = 重新定義 Loss function.

(step 2) → 原本的 Loss Function 考慮 error

改成 → 加上額外 term → $+ \lambda \sum (w_i)^2$
→ smaller are better.

→ smooth functions are better.

→ why? 較平滑的 function, 受 input 的影響較小

→ We prefer smooth, but don't be too smooth
(水平線, bad)