

# Introduction-draft

好的，這份 PDF 檔案中關於「新模型最佳化」的數學式整理如下，已使用 LaTeX 語法呈現：

## 1. 模型 (Model)

- 線性模型 (Linear Model)
  - $y = b + wx_1$  1
  - 其中， $y$  為觀看次數， $x_1$  為前一天的觀看次數， $b$  和  $w$  為未知參數（權重和偏差） 2。
- 新模型 (New Model)
  - 基於 Sigmoid 函數的模型： $y = b + \sum_i c_i \text{sigmoid}(b_i + \sum_j w_{ij} x_j)$  3
  - 基於 ReLU (Rectified Linear Unit) 函數的模型： $y = b + \sum_i c_i \max(0, b_i + \sum_j w_{ij} x_j)$  4

## 2. 損失函數 (Loss Function)

- 一般形式
  - 損失是參數的函數： $L(b, w)$  5
  - 誤差： $e = y - \hat{y}$  6
- 均方誤差 (Mean Square Error, MSE)
  - $L = \frac{1}{N} \sum_n (y_n - \hat{y}_n)^2$  7
- 平均絕對誤差 (Mean Absolute Error, MAE)
  - $L = \frac{1}{N} \sum_n |y_n - \hat{y}_n|$  8

## 3. 最佳化 (Optimization)

- 目標
  - 找到能使損失函數最小化的參數： $w^*, b^* = \arg\min_{w, b} L$  9
- 梯度下降法 (Gradient Descent)
  - 更新規則 (單一參數)： $w_1 \leftarrow w_0 - \eta \frac{\partial L}{\partial w} \mid w = w_0$  10
  - 更新規則 (多個參數)：
    - $w_1 \leftarrow w_0 - \eta \frac{\partial L}{\partial w} \mid w = w_0, b = b_0$  11
    - $b_1 \leftarrow b_0 - \eta \frac{\partial L}{\partial b} \mid w = w_0, b = b_0$  12
  - 向量化形式

- $\theta_1 \leftarrow \theta_0 - \eta g$  13
- 其中， $g$  為梯度向量： $g = \nabla L(\theta_0)$  14