

【機器學習2021】預測本頻道觀看人數 (上) - 機器學習基本概念簡介筆記

來源: 台灣大學教授 Hung-yi Lee

▶ 【機器學習2021】預測本頻道觀看人數 (上) - 機器學習基本概念簡介

Machine Learning

讓機器具備找一個函式的能力，也可以說是一個帶有未知parameters的function

Ex: $f(\text{一張圖片}) = \text{這是一隻熊貓}$

*Note:

$f()$: 是機器找出的函示，一張圖片: 是input，這是一隻熊貓: 是output

Machine Learning 不同工作類型 (根據output分類)

- ★ Regression
 - output: 是一個scalar
 - Ex: 預測Youtube隔天觀看人數
- ★ Classification
 - output: 根據給的options給出一個答案
 - Ex: AlphaGO
- ★ Structured
 - output: 給出一個有結構的物件ex: 文件、圖片

Training Steps



Step1:

Model: $y = f() = b + w * x_1$ (猜測)

Note: y : output， b (bias)、 w (weight)表示參數， x_1 : 已知資訊

Ex: 已Youtube觀看人數預測為例

y : 預測的今天觀看人數, x_1 : 前一天觀看次數, b 、 w : 從data中learn到的未知參數

Step2:

Loss: $L(b,w) =$ 這筆未知參數的數 (b,w) 是否好

將這筆參數的function帶入訓練的每筆data實際值，與預測值得到的誤差 (e) ，相加後得到的平均誤差值即是LOSS

最終LOSS值

Loss: $L = \frac{1}{N} \sum_n e_n$ ，表示參數的平均誤差

$$e = |y - \hat{y}| \quad L \text{ is mean absolute error (MAE)}$$

Step3:

Optimization: 找一個讓LOSS數值最小的w、b值

流程:

- Step1: 找到要走的方向
- Step2: 進行數值更新

$$\frac{\partial L}{\partial w} \Big|_{w=w^0, b=b^0}$$

$$w^1 \leftarrow w^0 - \eta \frac{\partial L}{\partial w} \Big|_{w=w^0, b=b^0}$$

$$\frac{\partial L}{\partial b} \Big|_{w=w^0, b=b^0}$$

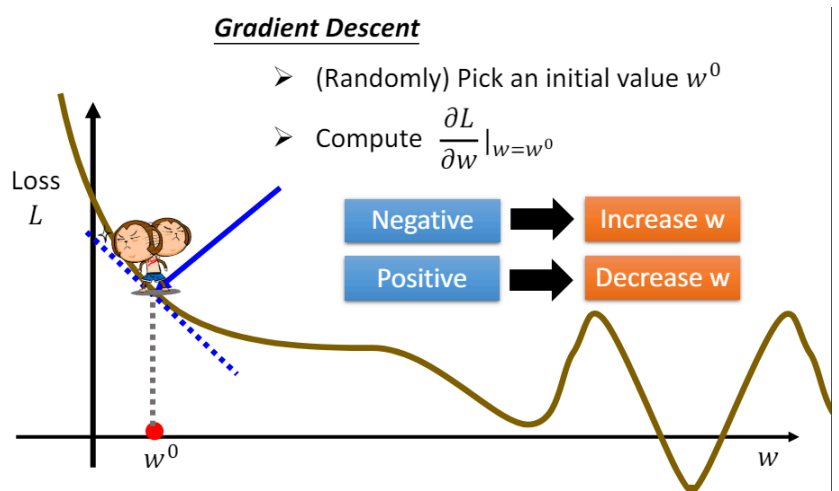
$$b^1 \leftarrow b^0 - \eta \frac{\partial L}{\partial b} \Big|_{w=w^0, b=b^0}$$

Step1

Step2

方法: **gradient descent**

計算斜率 (用微積分) · 判斷要往哪個方向走才能靠近最低點 · 若斜率是負的 · 則代表左高又低 · 理論上來說往右才能找到最低點; 正的則同理。

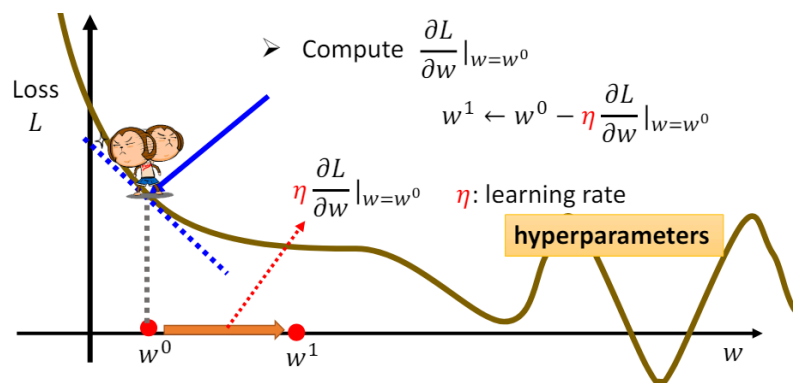


決定往哪裡走後 · 就是一次走多大的問題 · 這和learning rate有關。

hyperparameter定義: 代表需要自己設定的參數

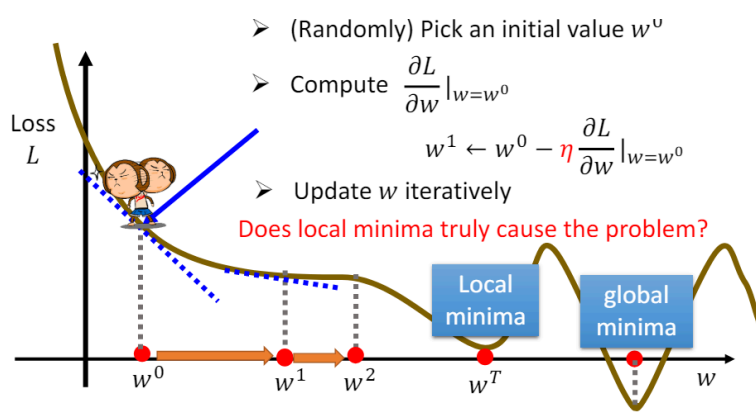
式子理解:

初始位置(w^0) - learning rate * 斜率
(斜率為負數則相加往右走; 為正同理)



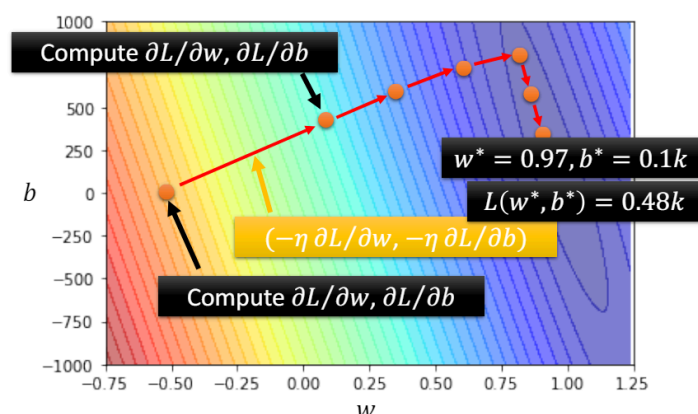
在gradient descent中有可能遇到Local minima的問題 (優化停在了不是正解的位置)

補充: 這裡老師說, 其實這不是真正的問題 😊 期待老師之後的說明!!



結論: 過程圖

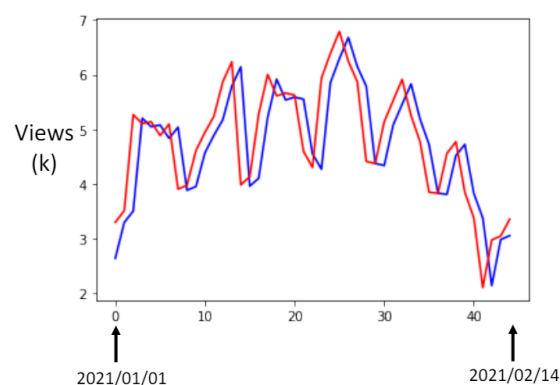
每個紅箭頭代表每一次的更新



最後老師用他一直舉的那個Youtube觀看人數預測例子, 繼續深入

從這張圖可以觀察到, 基本上觀看人數是以7天為一個循環在變動, 因此要改成讓model一次看7天的資料

這樣因為專業知識而改變Model, 可以說是 **Domain Knowledge**



改成7天後, 發現誤差有明顯下降。

其中各個 w_1^* , w_2^*下面的數值代表: 前一天、前兩天.....資料的重要性, 像在這之中, 胎人為前一天的資料最重要 ($w_1^* = 0.79$)

而老師在這中用的是 **Linear Model**:

($y = f() = b + w * x_1$)

| | | |
|-----------------------------------|----------------------------|----------------------|
| $y = b + w x_1$ | 2017 - 2020 $L = 0.48k$ | 2021 $L' = 0.58k$ |
| $y = b + \sum_{j=1}^7 w_j x_j$ | 2017 - 2020 $L = 0.38k$ | 2021 $L' = 0.49k$ |
| $y = b + \sum_{j=1}^{28} w_j x_j$ | 2017 - 2020 $L = 0.33k$ | 2021 $L' = 0.46k$ |
| $y = b + \sum_{j=1}^{56} w_j x_j$ | 2017 - 2020 $L = 0.32k$ | 2021 $L' = 0.46k$ |

Linear models

| b | w_1^* | w_2^* | w_3^* | w_4^* | w_5^* | w_6^* | w_7^* |
|-------|---------|---------|---------|---------|---------|---------|---------|
| 0.05k | 0.79 | -0.31 | 0.12 | -0.01 | -0.10 | 0.30 | 0.18 |

備註:

此筆記中的所有參考資料皆來自Hung-yi Lee 教授的上課影片和投影片，非常感謝Hung-yi Lee 教授的精彩課程，讓我重拾對人工智慧的興趣，不再覺得人工智慧像是一團迷霧~

此筆記目的是為了可以讓我更好的記得上課內容，並且做自己的學習紀錄~

若有侵權會停止分享這個筆記!