

110590007 資工系大三 白宸安

#### 一、Regression equation in basic part：

我在 basic part 中使用的 equation 是單純的線性， $Y = w_0 + w_1 * X$ ，其中 Y 是 sbp，而 X 是 dbp。

會選用這個 equation 是因為在資料分割和預處理後，把資料畫在圖上，發現他們有類似線性的關係。

#### 二、Variables you used in the advanced part：

我在 advanced part 中使用的 Variables 有包含 temperature、heartrate、resprate、o2sat，而透過把資料畫出來發現，其中 heartrate 和 sbp 有較強烈的線性關係，而 resprate 和 sbp 有些微的線性關係，另外兩個則沒有明顯線性關係，但有觀察到可能和一天之中的時間週期有些相關性。

使用的 equation 是  $Y = w_0 + w_1 * \text{temperature} + w_2 * \text{heartrate} + w_3 * \text{resprate} + w_4 * \text{o2sat}$ ，其中因為 temperature、o2sat 和 sbp 沒有明顯線性，所以這兩著的 weight 都被設為 0，在運算過程中也不去更新，所以實際只有  $w_0, w_2, w_3$  有係數更新。

另外，這個 model 是針對每個病人去算一組 weights，其中病人 14699420 和 15437705 這兩人的 heartrate 比較不是單純線性，因此這兩位的 heartrate 變數是使用平方，即他們的 kernel 是使用  $x^2$ ，其他人的則是  $x$ 。

#### 三、The difficulty you encountered

遇上的問題主要有：

1. 不太熟悉 np 的資料分割與預處理。
2. 把講義上的公式轉換為程式裡的式子時比較困難。
3. 分析資料的相關性

#### 四、How you solve the difficulty and your reflections

1. 開始做作業的前兩天花了不少時間來研究 np 操作，突破這個困難點之後，後續就進行的比較順利了。
2. 這裡轉換時有去網路上查一些相關作法，如 np 有現成的公式可以算 matrix inversion，gradient descent 也有蠻固定的流程。
3. 透過畫圖出來觀察資料，在做 model 選擇上才能更對症下藥，也比較有效地去降低 MAPE 值。

最後是心得，這次作業讓我體驗到很完整的機器學習流程，從一開始的分割資料和預處理，到實作模型函式，還有計算 MAPE 和畫圖分析等，經過這次作業，我對機器學習實作上更有概念。期待之後的作業跟 project。