# 國立台灣大學 物聯網導論期末專題報告

# 睡眠分析檢測解決方案

學生: 葉冠宏 R11943113

任文澤 R11942166

指導老師:淡江大學廖書漢教授

中華民國 111 年 12 月 15 日

## 目錄

目錄	I
一、 動機與目的	1
二、 使用的方法 & 流程	2
2.1 介紹 Ardiuno	2
2.2 介紹使用的模組	3
2.2.1 高聲度傳感器	3
2.2.2 小高聲度傳感器	4
2.2.3 MX9814	5
2.2.4 MicroSD 卡模組	5
2.3 模組的連接方式	6
2.3.1 環境音檢測	6
2.3.2 環境音錄製	8
三、 資料處理	9
3.1 介紹資料	9
3.2 資料的處理	11
3.2.1MFCC 特徵	11
3.2.2PCA 分析	14
四、 模型介紹 (LSTM)	15
4.1 使用模型介紹	15
4.2 模型具體參數介紹	15
4.3 模型的測試	16
五、資料的串流、儲存、預測,以及視覺化呈現	17
5.1 資料的串流	17
5.2 資料的儲存	17
5.3 視覺化呈現	18
5.4 資料的預測通知發送	18
六、結論	19
6.1 遇到的困難&解決&不足	19
6.2 總結	20
七、 参考文獻	
Λ <b>, </b>	21

## 一、動機與目的

睡眠質量顧名思義是用來判斷睡眠好壞的重要指標,它與身體健康情況息息相關,不好的睡眠質量會對人體健康產生影響,嚴重者甚至會危害生活品質,像睡眠呼吸中止症就是典型會破環睡眠質量的疾病,發病時會導致睡眠中的人產生雜亂的打鼾聲。而現代人隨著社會壓力的逐年遞增、以及越來越多的不健康生活習慣,導致部分人有著睡眠問題,但是若是想去醫院做睡眠質量的檢測,需要在相關的睡眠中心戴著檢測儀器進行睡眠測試,這會導致很多不習慣的人無法入睡或者其他心理因素導致睡眠檢測不准。出於對患者便利以及使結果更加準確的考慮,我們選擇使用睡眠時期的呼吸環境音以及深度學習的相關方法來對呼吸睡眠中止症的檢測,並且構建了一整套 ardiuno 錄音-即時回傳-睡眠質量檢測-訊息即時通知病發監護人的智慧型物聯網體系,使患者在日常居家睡眠中就能夠及時發現睡眠呼吸中止症,以及對後續的針對性治療提供更好的時機。

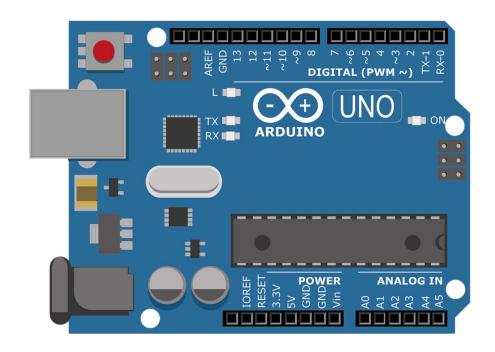


## 二、使用的方法&流程

睡眠呼吸中止症,是指在睡眠期間,呼吸暫停或者是呼吸減弱症狀導致的睡眠紊亂,每一次暫停期間可能從幾秒到數分鐘不等。在一般情況下,出現睡眠呼吸中止症的患者在睡夢中會產生吵雜的打鼾聲,或者是會產生窒息或者鼻音的出現。為了監測這一現象,我們使用 Ardiuno 對患者睡眠的環境音進行監測,同時使用 Ardiuno 對患者睡眠期間的環境音進行錄音,我們使用睡眠期間的環境音進行深度學習的相關分析,將標記好的睡眠的環境音進行 MFCC 特徵進行提取後,用 PCA 的方法降維,並放入 LSTM 神經網絡進行分析,若是出現睡眠呼吸中止症的情況,就使用 Line 即時提醒看護人員或者患者,使患者能夠即時得到相對應的治療。除此之外,我們也有把 Arduino 所接收到的資料存入 TimeScaleDB 的資料庫,並把即時的波動更新到我們所架設的網頁之中。至於物聯網中資料傳輸的部分,我們則有使用 MQTT 的架構來傳輸資料,以模擬真實世界的情況,以及解決讀寫速率不一、平行化運算等問題。

#### 2.1 介紹 Ardiuno

Ardiunos 是一個開源嵌入式的硬體平台,可以為使用者提供互動性的嵌入式專案。Ardiuno 有著很多不同模樣和不用功能的模組,可以通過不同的模組連接,從而實現各種各樣的炫酷的功能,比如說:空氣檢測裝置、Mini 四軸飛行器等。Ardiuno 也可以疊加其他板子從而實現更加多的功能。在我們的專題中,Ardiuno 負責的主要功能是錄音收集資料和實時傳輸訊息,其中用到了三個相關語音的模組。

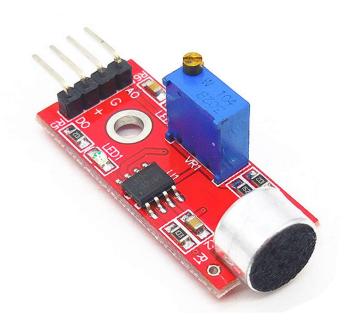


#### 2.2 介紹使用的模組

本次專題主要使用四個 Ardiuno 的相關模組從而分別實現檢測環境聲音、錄製環境音、存儲音檔的功能。這四個模組分別是高聲度傳感器、小麥克風、MX9814和 MicroSD 卡模組。

#### 2.2.1 高聲度傳感器

Big Sound 高感度聲音傳感器,高感度聲音傳感器對環境聲音的強度十分敏感, 一般用來檢測環境聲音的強度。當環境聲音沒有超過音量閥值。該模組會輸出高電平;但是環境聲音超過了設定的音量閥值的時候,會將麥克風聲音轉換為類比訊號 進行輸出。



#### 2.2.2 小-高聲度傳感器

小-高聲度傳感器類似上面的高聲度傳感器,唯一的區別是相比於高聲度傳感器,它有著一個更小的麥克風,一般也是用來檢測環境聲音的強度。當環境聲音沒有超過音量閥值。該模組會輸出高電平;但是環境聲音超過了設定的音量閥值的時候,會將麥克風聲音轉換為類比訊號進行輸出。通過調整音量的閥值,可以根據不同的閥值來出發環境聲音變化的條件。



#### 2.2.3 MX9814

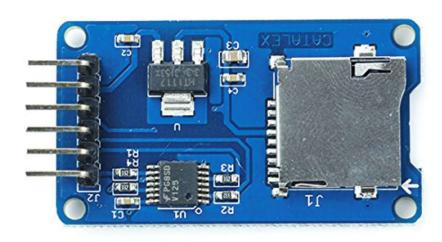
MX9814 是一個具有低雜訊自動增益控制的 MIC 放大模組,也就是麥克風放大器,有著自動增益控制(AGC)和低雜訊麥克風偏置。模組的低雜訊前端放大器,

可以增益放大(VGA),也有著輸出放大器、麥克風偏置電壓發生器和AGC控制電路。在本次專題中 MX9814 主要的功能是進行患者環境音的錄製,也就是資料收集的功能,將錄製好的語音檔案傳到 MicroSD 卡中進行存儲。



#### 2.2.4 MicroSD 卡模組

MicroSD Card Adapter 是 Mrcro SD 卡讀寫模組,通過檔案系統以及 SPI 介面的舉動程式,單機片系統即可完成 MicroSD 卡內的檔案讀寫工作。Ardiuno 用戶可以直接使用 Ardiuno IDE 自帶的 SD 卡程式庫即可完成卡的初始化和讀寫。在本次專題中,MicroSD 卡模組的主要功能是對患者的環境音的語音資料庫檔案進行存儲工作。

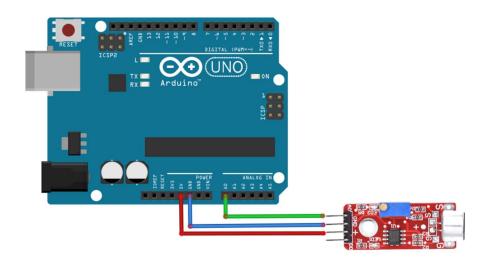


## 2.3 模組的連接方式

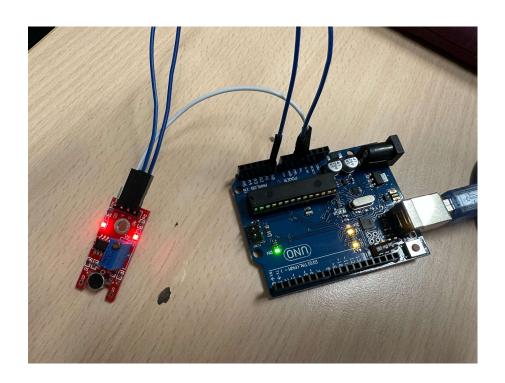
這部分會展示與 Ardiuno 相關的模組實現功能的電路模擬圖和實物連接。

### 2.3.1 環境音檢測

電路模擬圖:

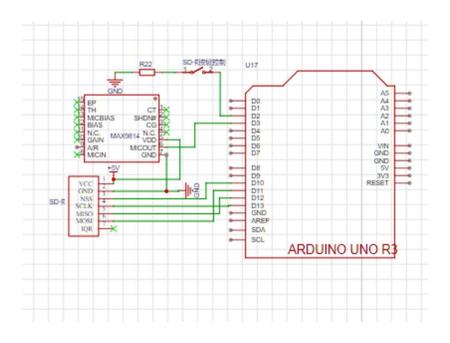


## 實物連結圖:



## 2.3.2 環境音錄製

## 電路模擬圖:

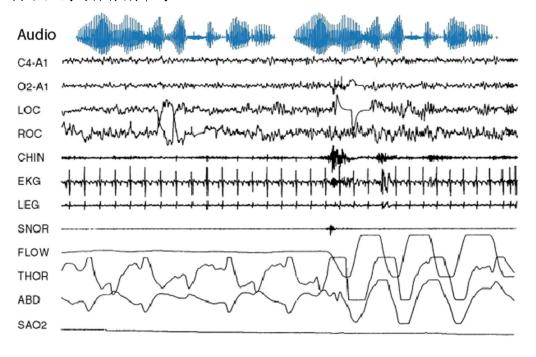


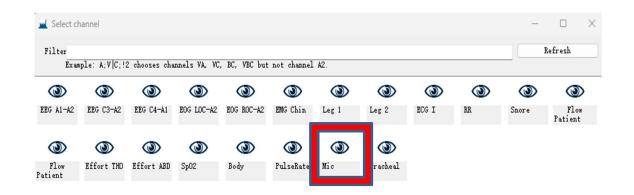
## 三、資料介紹以及處理

在訓練資料上面,我們把收集到的環境音音檔資料儲存完成之後,就進入語音資料的處理階段了,因為原始的資料是每一個 frame 下 normalized 過的數值, 我們必須做一些前處理,以讓我們的資料可以更好的去配適我們的模型。資料的前處理可以包括語音的轉換、特徵提取、特徵降維等一系列的操作。

#### 3.1 介紹資料

我們使用的訓練語音資料是 Sciecne Bank 的 PSG-Audio。 PSG (Polysomnography) 睡眠多項生理檢查,又被稱為多道睡眠紀錄,是睡眠醫學、睡眠障礙、打鼾、睡眠呼吸中止症等有關睡眠疾病中,最常被安排進行的標準生理診斷方法。多項生理檢查其中包括了多種儀器的配合進行的綜合診斷方法,比如其中包括腦波圖(EEG)、肌電圖(EMG)、心電圖(EMG)、喉嚨呼吸音(Tracheal)、環境呼吸音(Mic)...等。我們使用的 PSG-Audio 資料集包括了 212 位患者,每位患者大約錄製有 5 個小時的資料,以及其中包含了睡眠呼吸中止症的醫療團隊標記。其中他們的採樣頻率為 48000 Hz。





在資料集中,我們所使用的檔案主要包括 APNEA\_EDF 和 APNEA\_RML 兩個重要文件夾,其中 APNEA\_EDF 紀錄了患者的編號和睡眠多項數據的資料,而 APNEA\_RML 中紀錄了醫療專業人員所標記判斷睡眠過程中是否有發生呼吸中止症(ObstructiveApenea 和 Hypopnea),其在錄製第幾秒的時候發生,以及持續了幾秒。根據文獻,睡眠呼吸中止症一般要持續超過 10 秒才會被認定。



#### 3.2 資料的處理

患者的睡眠數據就是上節所展示的 EDF 檔案,它是集成了很多不同的通道的患者資料。我們曾經對各種頻道進行傅立葉分析比較有無睡眠呼吸中止症的差異,然而大部分均無明顯的差異,推測原因是雜訊太多等因素,僅有 Effort ABD 可以約略看出差別。然而我們的目的是希望使用睡眠時候的環境呼吸音來幫助患者解決睡眠呼吸中止症的判定問題,從而使患者將來檢查睡眠呼吸中止症的時候不需要在配戴過多的檢測儀器即可偵測,所以我們利用 PSG-Audio 資料集中的環境音

(Mic 通道)來進行我們的數據分析和訓練數據,這樣就可以使患者在家使用一個 簡單的錄音設備就能夠進行睡眠呼吸中止症的判斷。

除此之外,在原始資料提供的是每個 frame normalized 後的強度的數值,我們曾經嘗試拿每個 frame 的標記拿去訓練,也曾嘗試把完整 5 個小時的資料拿去做訓練,但可能由於資料不平衡的問題,效果並不好。因此我們僅拿標記為有症狀的前後一段時間作為訓練。而在測試的時候才去拿整段完整的資料,並以 window size 為 10 秒, stepsize 為 5 秒來做測試。

有關資料的處理,我們曾經做了以下嘗試:

1.預強調:將語音訊號通過一個高通濾波器,這個步驟主要是為了消除發聲過程中 聲帶和嘴唇的效應,來補償語音訊號受到發音系統所壓抑的高頻部分。

2.音框化:將N個採樣點集合成一個觀測單位,稱之為音框(Frame)。為了避免 相鄰的音框變化過大,兩個相鄰的音框會有一個重疊的區域。

3.快速傅立葉轉換:由於訊號在時域 (Time Domain) 上的變化通常很難看出訊號的特性,所以通常會將訊號轉換成頻域 (Frequency Domain) 上來觀察訊號的能量特徵分佈,每個音框 (Frame) 必須要經過快速傅立葉轉換的到訊號在頻譜上面的能量的分佈。

4.三角帶通濾波器:將快速傅立葉轉換後得到的能量頻譜乘以一組三角帶通濾波器, 得到每一個濾波器輸出的對數能量,這樣做的目的是將語音頻譜進行平滑化,並且 消除諧波,將原來語音的共振峰突出,同時也降低資料量。

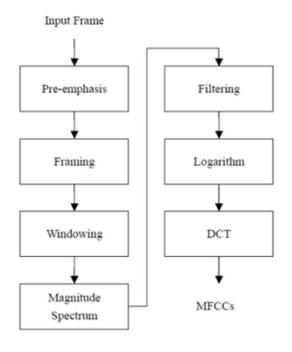
#### 3.2.1 MFCC 特徵

由於傅立葉分析並無提供特別的差異,因此我們借用了語音處理中 mfcc 的概念來對我們的資料進行轉換,作為特徵的提取。MFCC 特徵,也就是梅爾頻率倒譜係數(Mel-Frequency Cepstral Coefficients),是組成梅爾頻率倒譜的係數。梅爾頻率倒譜的頻帶劃分是在梅爾刻度上面劃分的,它相比於正常的對數倒頻譜中的線型間隔的頻帶更能近似人類的聽覺系統。

在 MFCC 中,我們取每單位 13 個維度來做轉換。然而由於資料中可能存在著雜訊,又每段標記資料的秒數並不相同,造成特徵的維度無法統一,因此我們把每個維度的平均數(可選用 hamming window)、標準差、維度 difference 的平均數、標準差作為特徵,因此每個 window 共可以提取到 13x4=52 個特徵資料。 MFCC 在特徵的提取可以分成以下幾個步驟:

1.對數能量:一個音框的能量是語音的重要特徵,我們通常還需要給語音的音框能量加上對數能量,我們取一個音框內的訊號平方和,再取以 10 為底的對數值,最後在乘以 10,使一個音框的語音特徵就有 13 維,其中是一個對數能量和 12 個倒頻譜參數。

2.差量倒數頻譜:求出 13 個特徵參數,實際應用時會加上差量倒頻譜參數,以顯示倒頻譜參數對時間的變化。



#### 3.2.2 PCA 分析

我們曾經把 52 維的 MFCC 資料拿去模型中做訓練,然而可能由於特徵的數量還是太多,容易造成 overfitting,因此效果並不佳。因此我們曾經嘗試了以下方式做資料的降維:

- 1. mask: 我們曾經嘗試 scikit-learn 的套件用 feature selection 的 tree 去選取其中幾項特徵,再針對原始 52 維的資料去做 mask 選取其中幾項作訓練,但效果並不佳。
- 2. LDA:Lda 是一種方法可以找到一種投影使得二元的目標函數中,data 可以 在組內的變異數越小越好,組間的變異數越大越好。但嘗試的結果是,效 果並不佳。
- 3. PCA 分析也就是主成分分析(Principal components analysis),這是一種統計分析、簡化數據機的方法。它利用正交轉換來對一系列相關變數的觀測值進行線性轉換,從而投影為一系列不相關變數的值,這些不相關的變數被稱之為主成分,主成分可以看作一個線性方程式。PCA 分析經常用於減少數據集的維數,同時保留數據集當中對變異數貢獻最大的特徵。這是通過保留低維主要成分,忽略高維主成分做到的,這樣低維成分往往能夠保留數據最終重要的成分。我們設定保留原始資料中 95%的變異數,通過 PCA分析,將 52 個維度的特徵轉換到 4 個維度的座標,這樣我們就能在沒有損失什麼資訊的情況下,用更簡潔的方式呈現該組數據,抓住原來維度數據的大部分特徵。

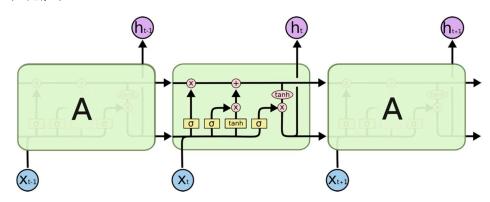
## 四、模型介紹

本章節將對我們使用的具體的睡眠呼吸中止症的預測模型進行介紹,經歷過了處理資料庫、特徵提取、特徵降維,我們就進入模型的訓練。我們曾經嘗試使用 Random forest 的架構去做模型的訓練。Random Forest 是一種透過表決的方式來形成的決策樹模型。然而可能由於 overfitting 等因素,效果並不好。因此我們就選擇嘗試使用 Bidirectional LSTM,利用其可以具有 memory 記憶前後特徵來加以訓練等效果,可以有效地學習到聲音前後段的關係,以達到更好的訓練效果。

在模型資料的標記上面,我們把資料中標記為 Obstructive Apnea 和 Hypopnea 等判定為睡眠呼吸中止症的秒數設定為 1,沒有呼吸中止症的秒數設定為 0。我們是把整段有睡眠呼吸中止症的秒數去做 mfcc 的轉換以及之後的特徵處理來做訓練。而為了解決資料不平衡的問題,我們僅取症狀發生前後各兩段,並長度和症狀發生一樣秒數的資料,來當作標記為 0 的訓練資料。而這樣也可以使 LSTM 模型去學習到症狀發生前後的特徵關係差異。

#### 4.1 使用模型介紹

LSTM (Long Short-Term Memory)模型,長短期記憶神經網絡也是一種時間循環神經網絡 (RNN),相比於普通的循環神經網絡,LSTM 更適合處理和預測時間序列中間隔和延遲非常長的重要事件,而普通的 RNN 在輸入的時間序列過長的時候會導致指數爆炸或者梯度消失的問題,從而難以捕捉長期時間的關聯。



#### 4.2 模型具體參數介紹

1.input\_size: 輸入特徵的特徵維數,即每一行輸入的元素個數。本次專題中 我們使用的四維的特徵。

2.hidden\_size: 為隱藏層狀態的維度數,即隱藏層節點的個數,該維數值是自定義的,我們使用 50 個 node。

3.num layer:LSTM 的堆疊層數,我們使用 2 層.

**4.bias:** 隱層狀態是否帶 bias,默認為 true。Bias 是偏置值或者偏移值。沒有偏置就是以 0 為中軸,或以 0 為起點。

5.batch\_first: 輸入輸出的數據第一維是否為 batch\_size, 默認值為 false。

**6.dropout:** 是否在除最後一個 RNN 層外的其他 RNN 層後面加 droupout 層。 我們設定為 0.2。

7.bidirectional: 是否為雙向 RNN,默認為 false。LSTM 可以根據數據輸入 從左到右推導結果,然後在用結果從右往左反推導,看結果和原因之間是否可 逆。

8.early stopping:為了避免掉 overfitting 的問題。

9.output layer:我們用的是 sigmoid function,以利於二元目標函數的判斷。 10.gradient descent optimizer: 我們用的是 Adam optimizer。

**11.loss function:**因為我們的問題是二元的問題,所以我們使用 binary cross entropy

#### 4.3 模型的測試

我們將準備的資料放入到模型進行訓練和預測,並把訓練參數做儲存。接著,我們使用另一位患者的 EDF 檔案來作為我們的 LSTM 的測試集。跟上面資料提取一樣的步驟,選擇另一位患者的 EDF 檔案中 Mic 的頻道,然後放入我們的模型,對比結果的預測與真實標籤的情況。並設定 window size 為 10 秒, stepsize 為 5 秒去做測試,以模擬真實情況中即時的資料輸入。

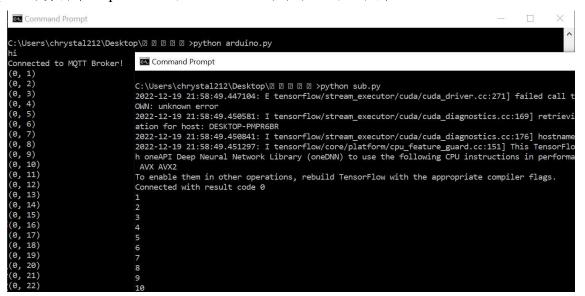
## 五、資料串流、儲存、預測,以及視覺化呈現

#### 5.1 資料的串流

MQTT 可以可靠地處理遞送訊息,並保持低成本地進行網路管理。 MQTT 提供了可高度擴充的發佈/訂閱通訊協定,且保證遞送。

在 arduino 接收到資料之後,我們去以 MQTT 的通訊協定去 publish 發送即時的資料到我們所設定的 topic 之中,而另一段則有 subscriber 去從一樣的topic 去做資料的抓取。因為後處理,例如:寫入資料庫、模型的預測...等需要花費比較長的運算時間,因此為了不耽誤讀資料的部分,所以我們才先把即時的資料以 MQTT 的方式先傳出,之後的資料儲存才交由平行化的方式去分配做運算。而在實際的物聯網應用中,我們也是應該把資料以網路協定的方式先行傳出,之後再從後端做更進一步的分析以及應用。

MQTT 的資料串流 publisher 和 subscriber 的部分呈現如下圖:

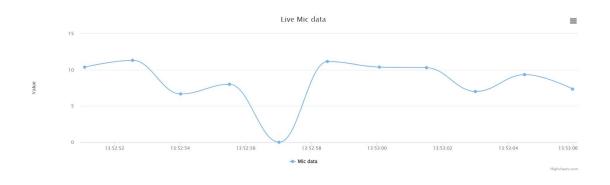


#### 5.2 資料的儲存

TimescaleDB 是一個開源的時間資料庫,可以使 user 快速的儲存和提取 資料。在 arduino 蒐集完每秒的資料後,我們去針對高於閾值的部分(為了不 讓背景雜音攤消掉有意義的音訊)去做平均,並把之存入 TimescaleDB 的時序 資料庫中。這部分我們是用平行運算的方式去做,因為寫入資料需要花費比 較多的時間。

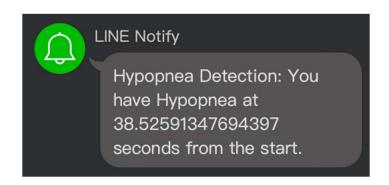
#### 5.3 視覺化呈現

針對存入資料的部分,我們會從資料庫中提取最新一秒的資訊,並更新到網頁中,以波動的方式作呈現。通過這個視覺化波形圖的呈現能夠讓看護人員更加直觀的了解患者睡眠呼吸中止症的情況。網頁的視覺化呈現如下圖:



#### 5.4 資料的預測通知發送

針對以 MQTT 所傳送的資料,我們會以 subscriber 去訂閱相同的 topic 去把原始的 data 作抓取,並略掉因為硬體的關係而產生的異常資料,累積一定的資料量後,傳送到我們所訓練的模型中做預測。如果確實有偵測到睡眠呼吸中止症,我們再以 Line Notify 的方式傳送通知以及偵測到的時間點 tag 到群組之中,希望能夠讓患者在第一時間得到對應的緩解方案或者治療。而值得注意的是,由於所 subscribe 的資料會比所 publish 的資料延遲許多,再加上模型的預測也需要時間,因此通知發送的時間點會比實際發生症狀的時間點還要再延遲個幾秒。不過這也是我們當初先以 MQTT 通訊協定傳出資料的原因,因為後處理需要時間,為了不要耽誤到讀取資料的進程。所呈現的通知如下圖:



#### 測試影片連結:

https://www.youtube.com/watch?app=desktop&v=jK4hTWGHUNQ

Demo 影片連結:

https://youtu.be/8fu25n0Z5WY

## 六、結論

#### 6.1 遇到的困難&解決&不足

- 1.資料處理的問題:在特徵的提取上面,我們曾經嘗試了許多的方法,也 針對不同頻道的數據進行分析,然而可能由於數據中參雜了許多的雜音, 因此有症狀和無症狀的部分有時候難以區分出差別。這部分我們花費了比較多的心力。
- 2.資料的選擇: PSG-Audio 給予的資料數量龐大,一共有將近 1TB 左右, 我們無法選擇所有的語音進行特徵篩選和訓練,所以在資料集的選擇上面 糾結了一些時間。而去我們本想使用喉嚨呼吸音進行特徵提取的(因為喉 嚨音更加明顯),而 Mic 呼吸環境音聲音偏小,但是為了我們的主題方便 患者不必使患者佩戴任何設備最終還是選擇了 Mic 通道的語音。
- 3.其實還有很多很好的 idea 但是沒有實現,因為 Mic 的環境呼吸音有著嘈雜的背景雜訊,想先使用其他神經網絡 CNN 或者 DeepFiliterNet 等網絡先進行雜訊的去除,再將頻譜放入 Transformer 網絡中進行睡眠中止症的辨識,但是時間有限,沒來的及完成。還有就是現在比較熱門的多模態方向,針對生物訊號加上一些其他圖片資訊進行神經網絡的訓練能夠大幅度提升預測結果的準確率。

#### 6.2 總結

很感謝淡江大學廖教授的幫助,給我們介紹了兩位正在負責相關領域的 學姐為我們講解整個專題,感謝台大和陽明交大的兩位學姐的幫助。

一套完整的 IOT 體系從提出到執行,需要考慮多方面的製作需求,需要兼顧整個物聯網體系的每一個環節,稍有錯誤就導致,滿盤皆輸。如何在適當的環境加入能夠方便使用者的技術都是我們必須考慮的內容。在經過一學期物聯網導論課程,我們對 IOT 整個大觀念和大的應用市場有著非常全面的了解,在物聯網背後時是各種各樣先進 AI 技術、通訊技術、電子技術的綜合。

在未來我們也將會讓我們上文提到來不及實作的 idea 進行嘗試實現,希望能夠有一天能夠真正的讓睡眠呼吸中止症的患者在家就能自我檢測,在家裡就能獲得準確的結果從而得到即時治療!

## 七、參考文獻

[1] Cho S-W, Jung SJ, Shin JH, Won T-B, Rhee C-S, Kim J-W. Evaluating prediction models of sleep apnea from smartphone-recorded sleep breathing sounds. JAMA Otolaryngol Head Neck Surg. Published online April 14, 2022. doi:10.1001/jamaoto.2022.0244

[2] Science Bank Dataset

https://www.scidb.cn/en/detail?dataSetId=778740145531650048

## 八、分工

葉冠宏:負責程式的撰寫、分析,ppt 的製作、書面報告的製作、Demo 影片製作任文澤:負責硬體的測試、ppt 的製作、書面報告的製作