

Machine Learning- Final Report

多模態病理嚔音分類

數據所 RE6124035 黃亮臻

Background

近年，隨著高壓的生活環境，各式各樣的文明病引起醫學界的注意，其中嗓音疾病常見於職業上需要大量用聲族群，如老師、業務、講師、零售攤販等，由於聲帶位處於喉部深處，喉部嗓音疾病檢測相當不易，需專業醫師操作特定儀器方能對病患進行診斷及治療，再加上現代人工作繁忙，不時有延誤就醫之情形。

在近年 COVID-19盛行期間，張口進行內視鏡檢查恐有飛沫傳播之風險。若能將人工智慧應用於非接觸式，**透過嗓音訊號(動態聲音)結合病史紀錄(靜態文字)偵測喉部病徵並分類**，有機會早期發現，早期治療，將會是所有嗓音患者之一大福音。

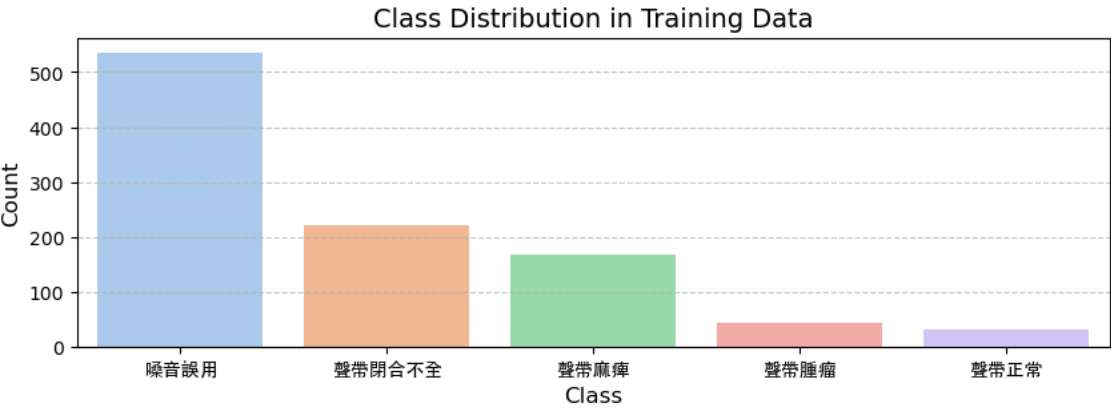
Data Description

- 資料來源:
 - 亞東醫院 - 耳鼻喉科嗓音資料庫
 - 教育部人工智慧競賽與標註資料蒐集計畫
- 資料格式: 病史紀錄(csv) + 嗓音訊號(wav)
- Train: 1000 位病人, Test: 500 位病人
- 病史紀錄: 26 個變數
- 嗓音訊號: 1~3秒鐘的 "啊"

病史紀錄

Y: 嗓音診斷分類

- 1. 嗓音誤用
- 2. 聲帶閉合不全
- 3. 聲帶麻痺
- 4. 聲帶腫瘤
- 5. 聲帶正常



| 變數 | Scale | 說明 |
|---------|---------|-------|
| 性別 | 1/2 | 男性/女性 |
| 年齡 | numbers | |
| 音域變窄 | 0/1 | 無/有 |
| 說話音量變小 | 0/1 | 無/有 |
| 說話久了容易累 | 0/1 | 無/有 |
| 喉嚨常覺得乾 | 0/1 | 無/有 |
| 喉嚨有異物感 | 0/1 | 無/有 |
| 胸口有灼熱感 | 0/1 | 無/有 |
| 吞東西容易噎到 | 0/1 | 無/有 |

病史紀錄

| 變數 | Scale | 說明 |
|-------|---------|--------------------|
| 眼睛乾澀 | 0/1 | 無/有 |
| 鼻涕倒流 | 0/1 | 無/有 |
| 糖尿病 | 0/1 | 無/有 |
| 高血壓 | 0/1 | 無/有 |
| 心臟病 | 0/1 | 無/有 |
| 頭頸部腫瘤 | 0/1 | 無/有 |
| 頭部損傷 | 0/1 | 無/有 |
| 腦中風 | 0/1 | 無/有 |
| 抽菸 | 0/1/2/3 | 從未/已戒菸/有抽菸/ 電子菸 |

| 變數 | Scale | 說明 |
|----------------------|-----------|--------------------------|
| PPD | numbers | 一天幾包菸 |
| 喝酒 | 0/1/2 | 從未喝酒/已戒酒/有喝酒 |
| 喝酒頻率 | 0/1/2/3 | 偶爾喝/每周喝/幾乎每天 |
| 症狀如何發生的 | 1/2/3/4/5 | 突然/逐漸變差/時好時壞/ 從小開始/其他 |
| 工作環境是否吵雜 | 1/2/3 | 否/有一點/很吵 |
| 聲音何時最差 | 1/2/3/4 | 早上/下午、晚上 /都一樣/ 不一定 |
| 用聲情形 | 1/2/3/4 | 總是需要/經常需要/偶而需 要/不需要 |
| VHI-10嗓音障礙指標 0 to 40 | | |

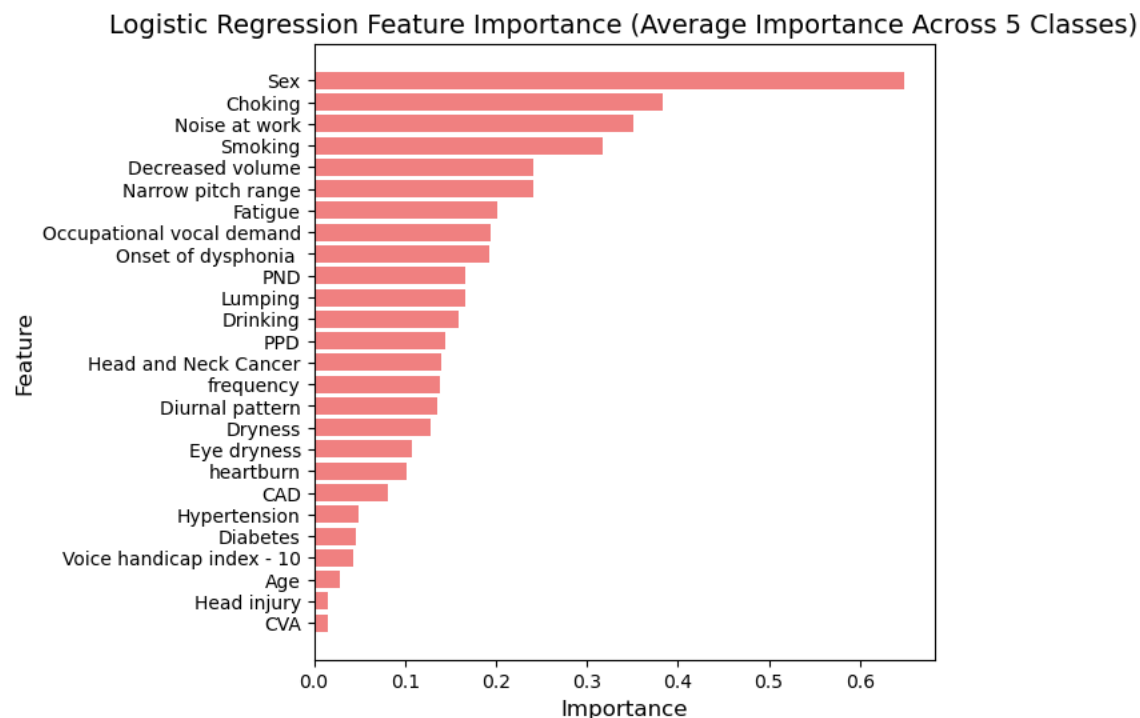
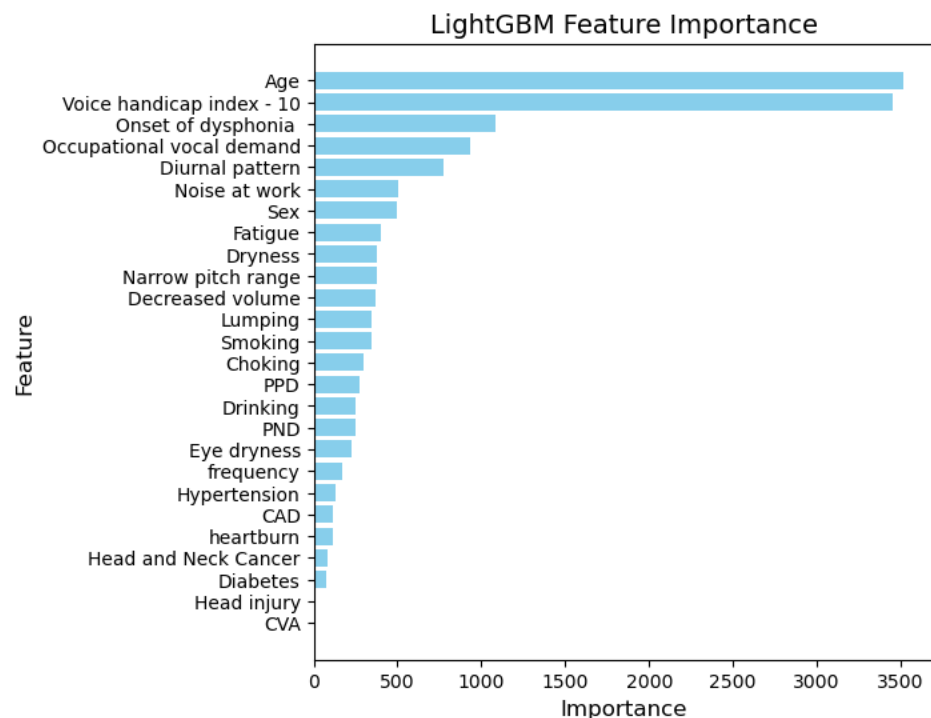
Missing Value

- PPD (一天幾包煙)：將從未抽過煙及已戒煙的人補0，其他取中位數
- VHI-10嗓音障礙指標：使用 KNN Imputer 補值

| Count of Missing Values | |
|---------------------------|-----|
| ID | 0 |
| Sex | 0 |
| Age | 0 |
| Disease category | 0 |
| Narrow pitch range | 0 |
| Decreased volume | 0 |
| Fatigue | 0 |
| Dryness | 0 |
| Lumping | 0 |
| heartburn | 0 |
| Choking | 0 |
| Eye dryness | 0 |
| PND | 0 |
| Smoking | 0 |
| PPD | 816 |
| Drinking | 0 |
| frequency | 0 |
| Diurnal pattern | 0 |
| Onset of dysphonia | 0 |
| Noise at work | 0 |
| Occupational vocal demand | 0 |
| Diabetes | 0 |
| Hypertension | 0 |
| CAD | 0 |
| Head and Neck Cancer | 0 |
| Head injury | 0 |
| CVA | 0 |
| Voice handicap index - 10 | 7 |

Feature Extraction

- 病史紀錄：LightGBM (Feature Importance) + Multinomial Logistic (Coef.)



最重要的前15個變數為：一天幾包菸, 抽菸, 音域變窄, 性別, 鼻涕倒流, 工作環境是否吵雜, 喉嚨常覺得乾, 說話久了容易累, 症狀如何發生的, 喝酒, 吞東西容易噎到, 聲音何時最差, 說話音量變小, 用聲情形, 喉嚨有異物感

Categorical Variable

Nominal 變數

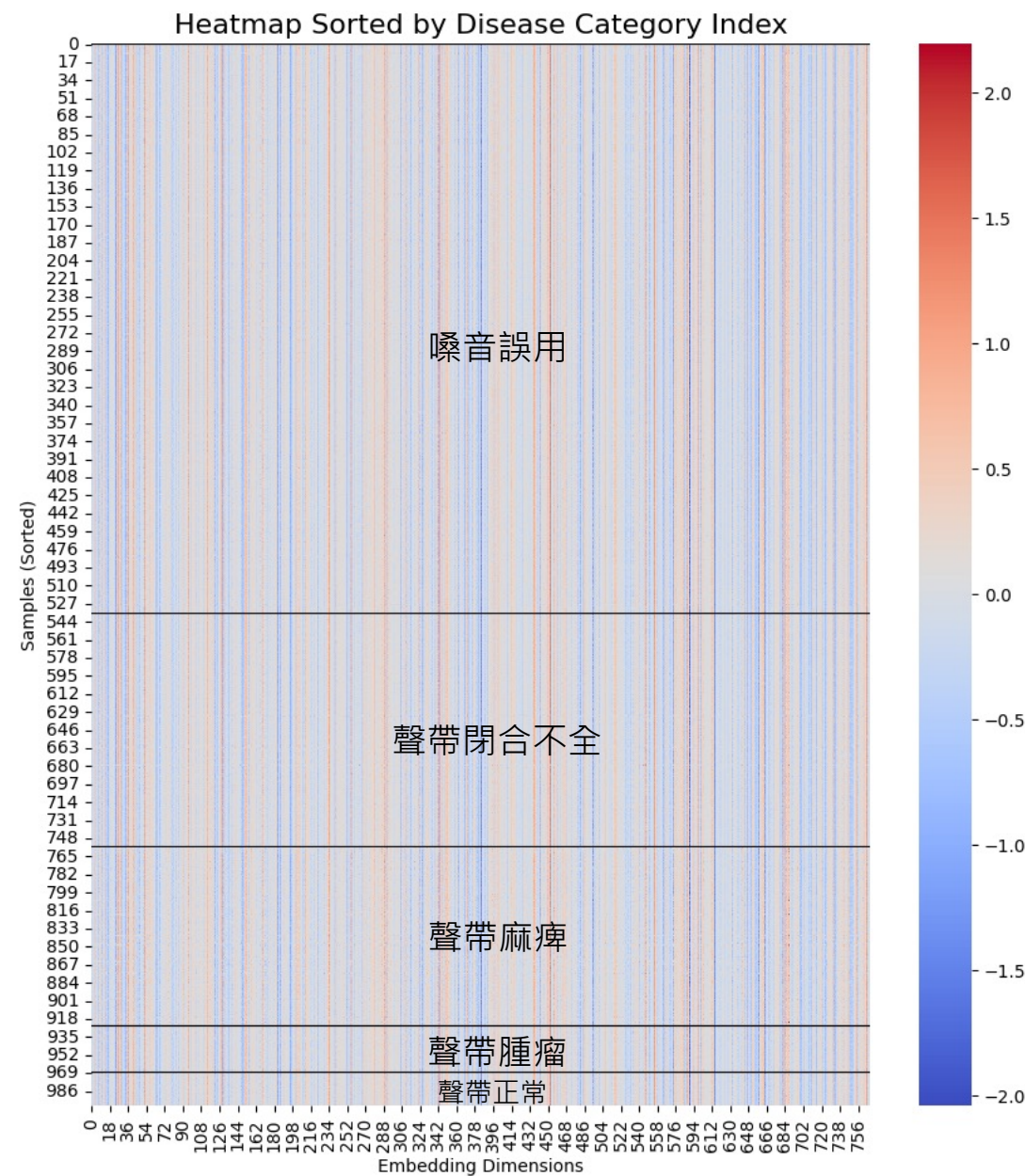
- 法一：~~One Hot Encoding~~ : Nominal 變數剛好都是兩類
- 法二：Multiple Correspondence Analysis (MCA)

Ordinal 變數：維持原樣

Ex: 抽菸 (1:從未/ 2:已戒菸/ 3:有抽菸/ 4:電子菸), 喝酒 (1:從未喝酒 /2:已戒酒/ 3:有喝酒), 用聲情形 (1:總是需要/2:經常需要/3:偶而需要/4:不需要)

嗓音訊號

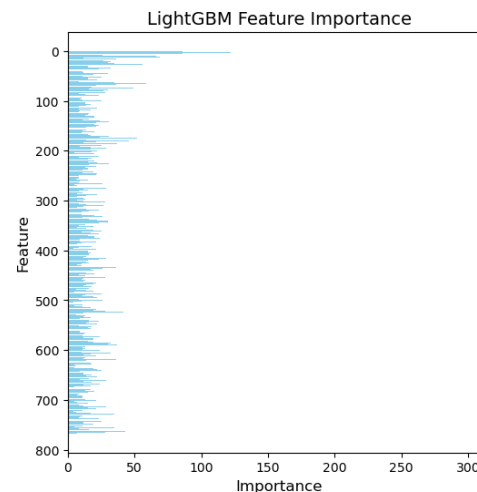
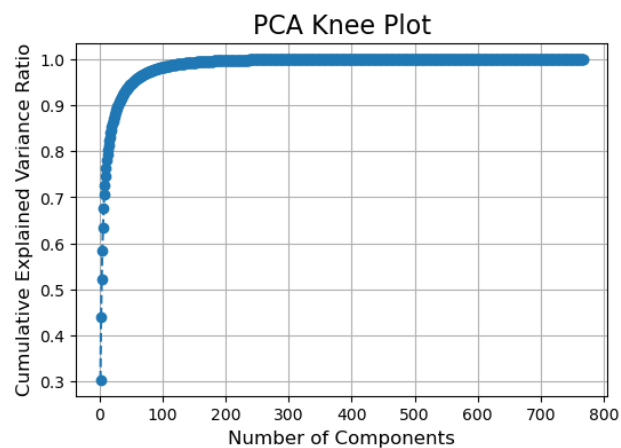
- 將音檔長度固定為2秒
 - 不足 2 秒：提取音檔的中間部分，複製並插入到音檔的中間，補齊 2 秒
 - 超過 2 秒：刪除中間段，確保音檔總長度縮減到 2 秒
- Fine-tune Wav2Vec2，得到 784 維 Embedding



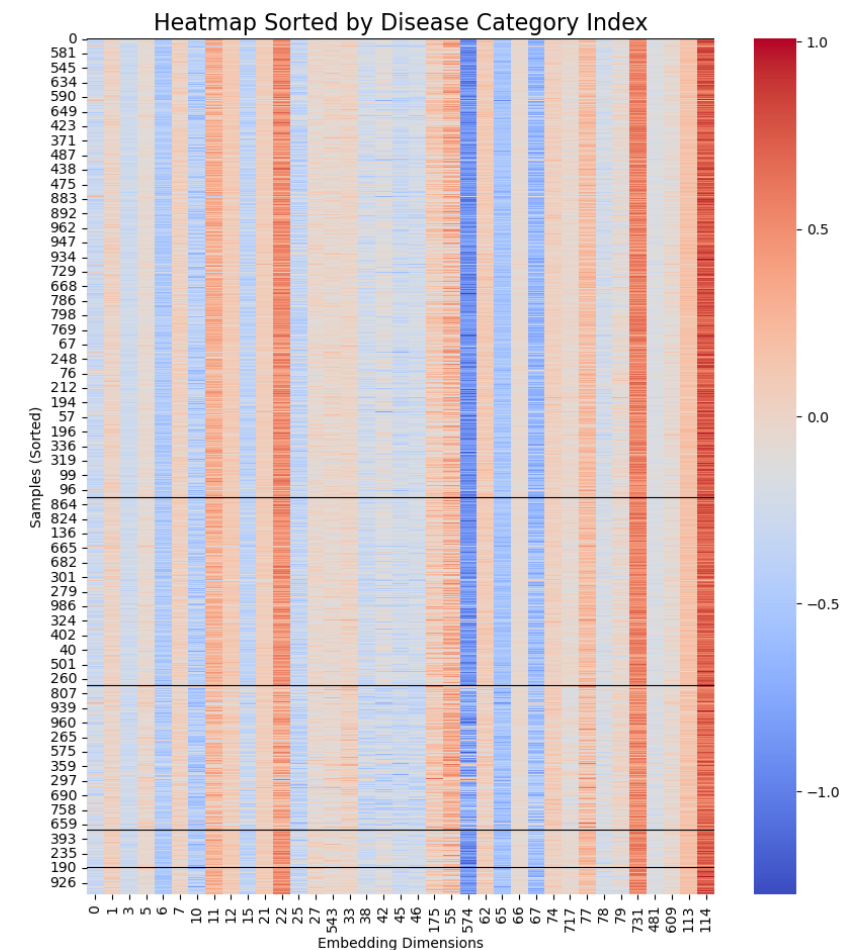
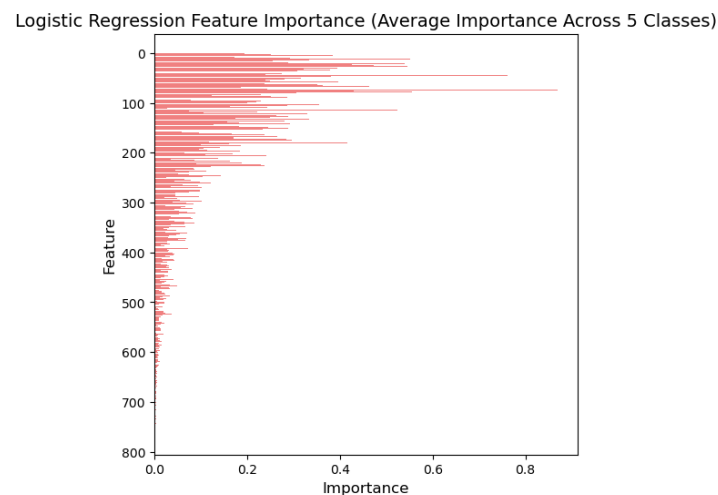
Feature Extraction

使用 PCA 將嗓音訊號濃縮特徵，
並將主成份再一次進行特徵選取

784維
➡



36個主成份
➡



實驗結果

| Method | Model | Accuracy | Recall (weighted) | Precision (weighted) | F1-Score (weighted) |
|---|----------|----------|-------------------|----------------------|---------------------|
| 病史資料(Feat. Sel.) | Logistic | 0.642 | 0.642 | 0.575 | 0.602 |
| | SVM | 0.694 | 0.694 | 0.631 | 0.654 |
| | XGB | 0.644 | 0.644 | 0.610 | 0.623 |
| | LGBM | 0.672 | 0.672 | 0.639 | 0.648 |
| 病史資料(Feat. Sel.) + 嗓音資料 PCA (Top) | Logistic | 0.728 | 0.728 | 0.688 | 0.701 |
| | SVM | 0.712 | 0.712 | 0.652 | 0.673 |
| | XGB | 0.698 | 0.698 | 0.658 | 0.669 |
| | LGBM | 0.700 | 0.700 | 0.654 | 0.666 |
| 病史資料(Feat. Sel.) + 嗓音資料 PCA (Feat. Sel.) | Logistic | 0.718 | 0.718 | 0.658 | 0.683 |
| | SVM | 0.714 | 0.714 | 0.657 | 0.677 |
| | XGB | 0.718 | 0.718 | 0.661 | 0.686 |
| | LGBM | 0.714 | 0.714 | 0.654 | 0.679 |
| 病史資料(Feature Selection) + 嗓音資料 PCA (Feat. Sel.) + MCA | Logistic | 0.738 | 0.738 | 0.675 | 0.701 |
| | SVM | 0.722 | 0.722 | 0.661 | 0.681 |
| | XGB | 0.724 | 0.724 | 0.663 | 0.689 |
| | LGBM | 0.715 | 0.715 | 0.655 | 0.655 |

Conclusion

- 透過將聲音特徵與病史資料結合，提升了模型的分類準確性。
- 病理嗓音的微小特徵容易被高維數據中的不相關訊息掩蓋。透過 PCA 的特徵濃縮與後續特徵選取，進一步挖掘真正有用的特徵。

Feature Extraction:

