洪徐永 r07525078

線性代數final project

鯨豚偵測演算法利用PCA分析

1. 研究動機

本實驗室為水下聲學實驗室，研究內容以水下音訊為主，包括了聲景研究、音傳路徑、目標強度等等，以及本次討論方向鯨豚偵測。鯨豚偵測又可細分為鯨豚偵測，鯨豚分類以及鯨豚定位，也就是DCL，detection-classification-localization。本研究欲透過所學以PCA，分析鯨豚哨叫聲頻譜圖形，透過被篩選出之主成分自動化辨識在此音訊時間段是否有鯨豚哨叫聲。

1. 文獻回顧

本研究參考了一篇線上文章，以及一篇論文。

A Step-By-Step Introduction to Principal Component Analysis (PCA) with Python ref:<https://datasciencesamurai.com/step-by-step-principal-component-analysis-pca-python/>

Feature Frequency Extraction Based on Principal Component Analysis and Its Application in Axis Orbit, Zhen Li , Weiguang Li , and Xuezhi Zhao

* 1. 線上文章 (A Step-By-Step introduction to Principal Component Analysis(PCA) with Python

作者以線上開源資料iris dataset 作為教材，從資料蒐集至資料正規化，至PCA分析，從頭教學PCA其中意義。其資料內容為三種不同品種的花，各包含4種特徵，如下表所示，最後利用這150筆線上開源資料，透過PCA分析，篩選其2重要的主成份，利用2D圖解釋其分群結果，如下圖所示。

|  |
| --- |
| 表 Iris dataset 特徵關係表 |
|  |

|  |
| --- |
|  |
| 圖 PCA 分群圖 |

* 1. 參考論文（Feature Frequency Extraction Based on Principal Component Analysis and Its Application in Axis Orbit）

此篇論文利用高速運傳的機器所傳遞出的振動時序列音訊資料，透過探討其有效的eigenvalue值與其頻域的關係，透過其震動聲音主成分判斷此機器是否保持在正常工作範圍。

1. 研究內容

本章節將詳述本演算法之資料前處理以及研究過程。本鯨豚哨叫演算法流程如圖 1所示。

首先將音訊資料輸入做短時距傅立業分析(STFT)，由於頻域解析度以及時域解析度成互補關係，在參數選擇上在不同的情況下會有不同的選擇，後續會做討論。之後對頻譜圖頻域軸做主成分分析，將頻譜資料從高維度降至低維度，篩選鯨豚哨聲其中的特徵。最後透過監督式學習將被降維之音訊頻譜資料分堆為訓練資料以及驗證資料，訓練最簡單之支持向量基模型(SVM)，最後以驗證資料驗證此模型之精確度。

|  |
| --- |
|  |
| 圖 1 鯨豚哨叫偵測演算法流程 |

* 1. 資料前處理與時頻分析

本研究使用之音訊原始資料為在台中港實海域所量測到之海豚哨叫聲，麥克風為底碇佈放，取樣頻率為96kHz，透過短時距傅立葉分析可透過人眼以及眼睛判斷是否有海豚哨叫聲，如圖 2所示。如圖海豚哨叫聲頻率範圍為3-8kHz，經過短時距傅立業分析可清楚辨識其窄頻以及高時間延續性之特性。

|  |
| --- |
|  |
| 圖 2 台中港音訊資料短時距傅立葉分析圖 |

資料蒐集的部分，首先以人工的方式篩選，透過人眼以及人耳，過濾原始資料留下有哨叫聲資料，以及當時台中港背景噪音資料作為無哨叫聲之資料，分類方式如表 1所示，以下僅以部分表格方式呈現，開始-結束時間為實際觀測數值。透過人工篩選進一步挑出有效之資料，進行傅立葉分析。

表 1 哨叫聲開始-結束時間表

|  |  |  |
| --- | --- | --- |
| 開始時間(s) | 結束時間(s) | 有無哨叫聲 |
| 31.3 | 31.6 | 1 |
| 33.3 | 33.6 | 1 |
| 35.45 | 35.55 | 1 |

…

…

…

|  |  |  |
| --- | --- | --- |
| 250 | 251 | 0 |
| 256 | 257 | 0 |
| 284 | 285 | 0 |

短時距傅立業分析為最常見之時頻分析方法，數學公式如公式 1所示。M為窗函數長度，為音訊資料取樣頻率，而j為虛數，窗函數使用漢寧窗函數，M長度選擇為1024點，以90%重疊，時間解析度約為1ms，頻域解析度約為94Hz。

|  |  |
| --- | --- |
|  | ……………………….公式 1 |

* 1. 主成分分析

主成分分析可視為將原始資料特徵點以最佳的權重做線性組合，已形成分散程度高的資料點，降低各個特徵之間的相關性，以達到各個被保留下的主成分擁有高的資料密度，達到將原始資料降維的目的。

本研究將短時距傅立葉分析後之頻域軸能量作為原始資料。根據奈奎斯特取樣定理，由於音訊資料取樣率為96kHz，有效頻率由0Hz至48kHz，窗函數M長度為1024點，故原始資料維度為513維(N//2+1)。

實際上透過主成分分析後提取的主成分與原始資料的主成分維度相同，也就是說假設原始資料維度為10，經過主成分分析後其維度仍然保持10個。然而在大多數的分析裡只有少數特徵保有大量有意義的資料，將這些主成分保留下做後續的演算法使用將減少資料維度且提高演算法效率以及準確度。因此，決定哪些重要的成分需要被保留下來，哪些多餘的資料需要被拋棄就至關重要，本研究採用scree-test決定最後被保留下的主成分，如圖 3所示。本研究選取了20個有相對較大特徵值(eigenvalue)之特徵向量做為線性組合之權重，其變異數總和佔整體87%。也就是說，透過PCA我們將原始資料從維度513維降維至20維，但我們能保有原始資料的87%。

|  |
| --- |
|  |
| 圖 3 scree test 圖 |

* 1. SVM模型

在機器學習中，支持向量機(support vector machine)，常簡稱SVM，是在分類與迴歸分析中分析資料的監督式學習模型。我們將資料分為訓練資料，以及驗證資料，比重為8:2，此分類原因是為避免機器在學習過程中從驗證資料裡偷偷事先知道了答案，造成後續驗證階段的不準確;同時此分類方法可提供後續判斷機器是否有過度擬和原始資料的情況，機器過度擬和雖可在訓練資料達到極高的分類成功率，但在實際預測上將會有極差的表現。

本研究有1200筆資料，600比為鯨豚哨叫聲資料，600比為無哨叫聲當地聲景資料，其分組方式如表 2所示，將960筆資料作為訓練資料，在模型訓練階段，將模型與240筆驗證資料完全區隔開來，供後續作為對此訓練模型的評估。

表 2 訓練資料-驗證資料分組方式表

|  |  |  |  |
| --- | --- | --- | --- |
|  | 訓練資料 | 驗證資料 | 總數 |
| 哨叫聲資料點數 | 474 | 126 | 600 |
| 無哨叫聲資料點數 | 486 | 114 | 600 |
| 總數 | 960 | 240 | 1200 |

1. 結果與討論
   1. SVM 模型預測

將事先留下的240個驗證資料輸入已訓練好之SVM模型，表 3為其辨識結果之混淆矩陣，以此矩陣整理計算其相對應之準確率以及召回率如表 4所示。根據此表我們可以發現其整體準確率達93%，其結果是相當優秀的。此演算法有其中兩項指標須特別留意，其一為有哨叫聲之準確率，其數值為92%，也就是說，在100聲之確定有哨叫聲之音訊資料，有其中的92聲可被此模型偵測出來;另一有趣的參數為哨叫聲之召回率，其數值為94%，也就是說，在100聲模型聲稱有哨叫聲之音訊資料，其中的94聲為正確有哨叫聲之音訊資料，其誤報率為6% ，誤報率相當的低。

表 3 SVM模型辨識結果表

|  |  |  |  |
| --- | --- | --- | --- |
|  | 無哨叫聲(辨識結果) | 有哨叫聲(辨識結果) | 總數 |
| 無哨叫聲 | 103 | 11 | 114 |
| 有哨叫聲 | 7 | 119 | 126 |

表 4 準確率以及召回率表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | precision | Recall | F1-score | 總數 |
| 無哨叫聲 | 0.94 | 0.90 | 0.92 | 114 |
| 有哨叫聲 | 0.92 | 0.94 | 0.93 | 126 |
| 準確率 |  |  | 0.93 | 240 |

* 1. 未來工作

事實上，市面上已經有相當多鯨豚偵測演算法，本演算法雖已有高達93%的精確率，但演算法不僅僅只考慮其精準度，執行速度也相當重要，其重要程度攸關此演算法之使用時機，未來須再比較各個演算法之精準度以及執行速度;同時可再根據混淆矩陣之結果，回頭檢視是哪一類之音訊資料造成誤判，精進本鯨豚偵測演算法之精準度。