**NYCU 大學部人工智慧專題 - 專案 #1 報告**

**專案名稱：** 搖滾音樂分類器  
**姓名：** 林辰翰  
**學號：** 111550138  
**GitHub 數據集連結：** [填入你的 GitHub 連結]

1. **研究問題概述**

簡要描述你的研究問題，為何選擇這個主題，以及它的重要性。

1. **Dataset**
   1. **數據集概述**

* **數據類型**：mp3
* **數據來源**：Spotify上的歌曲
* **數據集大小與組成**：總共約200首歌曲，有以下5個類別

|  |  |
| --- | --- |
| 類別 | 數量 |
| Hard Rock | 32 |
| Metal | 57 |
| Pop Rock | 40 |
| Post Rock | 17 |
| Punk Rock | 64 |

* 1. **數據收集過程**
* **採集方法**：從spotify上複製歌曲或專輯連結，並利用<https://spotidownloader.com/> 來下載音樂
* **數據示例**：

Dataset\Metal\Wherever I May Roam - Remastered 2021.mp3

Dataset\Punk Rock\The American Dream Is Killing Me.mp3

1. **Methods**
   1. **Data Preprocessing**

* **Feature Extractor**: MFCC (librosa)
* **Dimensionality reduction**: PCA (sklearn.decomposition)
* **Encode Labels & Standardize the features**: sklearn.preprocessing
  1. **Supervised Learning**
* **方法 1**：Support vector machine (sklearn.svm)
* **方法 2**：Random forest (sklearn.ensemble)
  1. **Unsupervised Learning**
* **方法**：K-means (sklearn.cluster)

1. **Evaluation**

關於supervised learning的部分:

* 下列結果都是經由 cross-validation 得出，利用了sklearn.model\_selection的cross\_val\_predict
* 下列 metrics 是利用sklearn.metrics的classification\_report得出
* 這個結果是利用了全部的dataset
* 在data preprocessing的部分是將data的維度降到了10
  1. **Support vector machine**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| **Hard Rock** | 0.80 | 0.88 | 0.84 | 32 |
| **Metal** | 0.72 | 0.74 | 0.73 | 57 |
| **Pop Rock** | 0.63 | 0.65 | 0.64 | 40 |
| **Post Rock** | 0.85 | 0.65 | 0.73 | 17 |
| **Punk Rock** | 0.63 | 0.62 | 0.63 | 64 |
|  | | | | |
| **accuracy** |  |  | 0.70 | 210 |
| **Macro avg** | 0.73 | 0.71 | 0.71 | 210 |
| **Weighted avg** | 0.70 | 0.70 | 0.70 | 210 |

* 1. **Random forest**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| **Hard Rock** | 0.76 | 0.81 | 0.79 | 32 |
| **Metal** | 0.78 | 0.81 | 0.79 | 57 |
| **Pop Rock** | 0.64 | 0.70 | 0.67 | 40 |
| **Post Rock** | 0.67 | 0.47 | 0.55 | 17 |
| **Punk Rock** | 0.62 | 0.59 | 0.61 | 64 |
|  | | | | |
| **accuracy** |  |  | 0.70 | 210 |
| **Macro avg** | 0.69 | 0.68 | 0.68 | 210 |
| **Weighted avg** | 0.69 | 0.70 | 0.69 | 210 |

* 1. **K-means**
* **Adjusted Rand Index**: 0.15
* 有利用label來計算ARI
* ARI是由sklearn.metrics的adjusted\_rand\_score得出

1. **Experiments:** 對於SVM model所做的實驗
   1. **The amount of training data**

* **Use 50% of dataset**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| **Hard Rock** | 0.77 | 0.77 | 0.77 | 13 |
| **Metal** | 0.64 | 0.62 | 0.63 | 26 |
| **Pop Rock** | 0.62 | 0.60 | 0.61 | 25 |
| **Post Rock** | 1.00 | 0.44 | 0.62 | 9 |
| **Punk Rock** | 0.54 | 0.66 | 0.59 | 32 |
|  | | | | |
| **accuracy** |  |  | 0.63 | 105 |
| **Macro avg** | 0.71 | 0.62 | 0.64 | 105 |
| **Weighted avg** | 0.65 | 0.63 | 0.63 | 105 |

* **Use 80% of dataset**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| **Hard Rock** | 0.80 | 0.89 | 0.84 | 27 |
| **Metal** | 0.77 | 0.85 | 0.81 | 47 |
| **Pop Rock** | 0.60 | 0.60 | 0.60 | 25 |
| **Post Rock** | 1.00 | 0.62 | 0.77 | 16 |
| **Punk Rock** | 0.67 | 0.64 | 0.65 | 53 |
|  | | | | |
| **accuracy** |  |  | 0.70 | 168 |
| **Macro avg** | 0.77 | 0.72 | 0.68 | 168 |
| **Weighted avg** | 0.74 | 0.73 | 0.69 | 168 |

* **Use 100% of dataset**

請見4.1的表格

* 1. **The dimensions of the features**

在extract feature時，每首歌會取40個mfcc feature，接下來會探討將feature 利用pca減少到不同維度的影響

* **dim = 2**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| **Hard Rock** | 0.42 | 0.16 | 0.23 | 32 |
| **Metal** | 0.81 | 0.68 | 0.74 | 57 |
| **Pop Rock** | 0.55 | 0.70 | 0.62 | 40 |
| **Post Rock** | 0.47 | 0.41 | 0.44 | 17 |
| **Punk Rock** | 0.48 | 0.62 | 0.54 | 64 |
|  | | | | |
| **accuracy** |  |  | 0.57 | 210 |
| **Macro avg** | 0.54 | 0.52 | 0.51 | 210 |
| **Weighted avg** | 0.57 | 0.57 | 0.55 | 210 |

* **dim = 5**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| **Hard Rock** | 0.70 | 0.81 | 0.75 | 32 |
| **Metal** | 0.66 | 0.68 | 0.67 | 57 |
| **Pop Rock** | 0.52 | 0.55 | 0.54 | 40 |
| **Post Rock** | 0.69 | 0.65 | 0.68 | 17 |
| **Punk Rock** | 0.68 | 0.59 | 0.63 | 64 |
|  | | | | |
| **accuracy** |  |  | 0.65 | 210 |
| **Macro avg** | 0.65 | 0.66 | 0.65 | 210 |
| **Weighted avg** | 0.65 | 0.65 | 0.63 | 210 |

* dim = 10

請見4.1的表格

1. **Discussions**

* **結果是否符合預期？**

由於音樂分類是一個較具挑戰性的任務，模型的表現整體符合預期。SVM 和 Random Forest 在數據較少時表現一般，但當數據集增大時，準確率達到了 70% 左右。這顯示了機器學習模型對於音樂分類的可行性，但仍有提升空間。

* **數據集的哪些特徵影響了模型表現？**
  + **MFCC 特徵**：這是音訊處理中常用的特徵，對分類結果影響較大。
  + **維度降至 10**：經過 PCA 降維，保留 10 個主要特徵時，分類結果較為穩定，而當維度降至 2 或 5 時，準確度下降，顯示出部分關鍵信息可能被丟失。
  + **不同類別數據分布**：某些類別（如 Post Rock）樣本數較少，導致模型較難準確分類這些類別。
* **如果有更多時間，會進行哪些額外實驗？**
  + **增加數據量**：目前數據集僅有 200 首歌，擴充數據集可能能提高模型的泛化能力。
  + **嘗試其他特徵提取方法**：例如 Chroma Features 或 Spectral Contrast，可能有助於區分音樂風格。
  + **嘗試深度學習方法**：如 CNN 或 LSTM，可能能更有效地學習音樂的時序特徵。
  + **嘗試更好的聚類方法**：目前 K-means 取得的 ARI 僅 0.15，可能可以使用 DBSCAN 或 GMM 來改善無監督學習的表現。
* **從實驗中學到了哪些經驗？還有什麼未解決的問題？**
  + 音樂分類的難點在於類別間的特徵可能重疊，例如 Pop Rock 和 Punk Rock 的聲音特徵可能相似，使得分類結果不夠明確。
  + 樣本數量對結果影響很大，特別是少數類別的準確度較低，未來可以嘗試數據增強（Data Augmentation）。
  + 目前的分類方法基於統計特徵，但音樂的風格有時涉及更高層次的語義信息，可能需要更複雜的特徵提取技術。

1. **Conclusion**

* **總結研究結果與發現**
  + SVM 和 Random Forest 在使用完整數據集時，準確率約為 70%，表現相當穩定。
  + 維度降至 10 後效果最佳，證明 PCA 能有效降低冗餘特徵，提升分類表現。
  + K-means 在無監督學習下的效果不佳（ARI = 0.15），說明傳統聚類方法可能不適合音樂分類。
* **未來改進方向**
  + **擴充數據集**：增加歌曲數量，尤其是 Post Rock 類別，以平衡數據分佈。
  + **改進特徵提取**：嘗試 Chroma Features、Spectral Features，甚至 CNN 來學習更深層次的音樂特徵。
  + **探索深度學習方法**：如使用卷積神經網絡（CNN）來學習頻譜圖特徵，或 LSTM 來捕捉音樂的時序信息。
  + **改進聚類方法**：嘗試其他聚類技術，如 DBSCAN 或 GMM，提高無監督學習的分類準確度。

**附錄：程式碼**