胡志堅、陳昱安 (2024),「運用深度學習與主題模型建構歌曲風格和歌詞意涵之整合分析機制」,*資訊管理學報*,第三十一卷,第二期,頁 209-237。

運用深度學習與主題模型建構歌曲風格和歌詞意涵之整合分析機制

胡志堅 *
大同大學資訊經營學系

陳昱安
大同大學資訊經營學系

摘要

資通訊科技的蓬勃發展,驅使眾多音樂愛好者透過音樂串流服務平台聆聽和分享音樂創作。然而,隨著音樂作品數量的增加,有效管理這些作品並提升音樂檢索效能,成為音樂數位典藏的重要課題。目前的音樂素材檢索和歸類,缺乏同時考量音樂風格及其創作之文化背景。因此,本研究嘗試運用深度學習、以及文字探勘等資訊科技,來分析歌曲風格和歌詞之間的關聯。首先,採用卷積神經網路模型(CNN)進行歌曲風格分類,接著結合組合式主題模型(CombinedTM)分析歌詞主題傾向。研究結果顯示,所建立的歌曲風格分類模型能準確分類音樂風格並解析其音樂的組成成份。透過歌詞主題傾向雷達圖分析歌詞內容,也能解釋不同類別歌詞的意涵。整合這兩種機制能更深入地檢索音樂作品,並連接歌曲風格與歌詞意境,以及相關背景知識。建議將來可將歌曲風格音樂成分分析與歌詞主題傾向雷達圖分析整合到音樂數位典藏系統和網路音樂串流平台,提升音樂作品的檢索效能,並建立音樂知識脈絡。

關鍵詞:音樂風格分類、卷積神經網路、音樂資訊檢索、文字探勘、主題模型

* 本文通訊作者。電子郵件信箱:holdenhu@gm.ttu.edu.tw
2023/07/09 投稿;2023/09/17 修訂;2023/10/20 接受

Hu, C.C. & Chen, Y.A. (2024). Constructing an Integrated Analysis Mechanism for Music Genre and Lyric Semantics using Deep Learning and Topic Modeling. *Journal of Information Management*, 31(2), 209-237.

Constructing an Integrated Analysis Mechanism for Music Genre and Lyric Semantics using Deep Learning and Topic Modeling

Chih-Chien Hu*
Department of Information Management, Tatung University

Yu-An Chen
Department of Information Management, Tatung University

Abstract

The development of information and communication technology has led to many music enthusiasts using music streaming platforms to enjoy and share music. However, with the increasing number of music works, effectively managing these works and improving music retrieval efficiency have become important issues in digital music preservation. The current classification and analysis of music materials often overlook the music genres and cultural backgrounds. Therefore, this study utilizes information technology to analyze the relationship between songs and lyrics. Firstly, a convolutional neural network (CNN) is used for song genre classification, followed by the combination of a composite topic model (CombindTM) to analyze the thematic tendencies of lyrics. The research results show that the established song genre classification model can accurately classify music genres and analyze the compositional elements of music styles. By analyzing the lyrics' thematic tendencies using a radar chart, the textual meaning of lyrics in different categories can also be interpreted. Integrating these two mechanisms allows for a deeper retrieval of music works and the synthesis of song styles, lyrics' moods, and relevant knowledge. It is suggested to integrate both mechanisms into digital music preservation systems and online music streaming platforms in the future to enhance the retrieval efficiency of music works and establish the context of music knowledge.

Keywords: Music genre classification, Convolutional Neural Network (CNN), Music Information Retrieval (MIR), Text mining, Topic model

* Corresponding author. Email: holdenhu@gm.ttu.edu.tw

2023/07/09 received; 2023/09/17 revised; 2023/10/20 accepted

壹、緒論

隨著網絡音樂串流服務平台不斷擴張，吸引了眾多音樂愛好者藉此平台聆聽、分享，甚至投入網絡音樂創作。截至 2021 年，全球網絡音樂串流訂閱服務的用戶數量已經突破 5.24 億人(Mulligan 2022)，增長速度驚人，已經成為流行音樂發展的重要媒介。然而，隨著平台上音樂作品數量的不斷增加，如何有效管理這些音樂作品，提升音樂檢索的效能，並提供有意義的音樂知識給用戶儼然是音樂數位典藏領域一個值得探討的問題。

音樂的組成元素通常包含節奏(rhythm)、旋律(melody)、和弦(chord)、和聲(harmonic)、音色(timbre)、以及力度(dynamics)等(Byrd & Crawford 2002; Downie 2003)。節奏規範著樂曲的規律性，將音樂區分成許多部份，並產出律動。旋律是藉由連續的音符和節奏(如拍子、速度)所組成，透過旋律的組成形式能夠提升樂曲情境的豐富性。和弦的基本組成為三和弦（triad），為了讓音樂更有變化性，在和弦上添加不同的音符即能組合成新的和弦，和弦的產生可以增加旋律的多樣變化輔助旋律的運行。和聲的基本單位為和絃，透過和聲理論可以建構大調(major)、以及小調(minor)等樂性，以陪襯旋律創造出多變的效果(Schmuckler 1989)。音色是指各種樂器(或人聲)因發聲體本身材質或形狀受到作用時所呈現的特色聲響，在音樂創作時樂器音色的選擇亦相當重要。力度是一種表示聲音強、弱程度的相對概念，藉由力度標記得以顯示聲量從最弱到最強之狀態變化，每一個層次的力度都是一個相對值。綜合各種音樂元素的調整與變化，即可創作出多元風格的音樂(Byrd & Crawford 2002)。

一首完整的歌必然結合歌曲與歌詞，其不僅具有娛樂性，同時也呈現時代文化的軌跡。歌詞是追隨著歌曲的結構文字，一首歌如同一個故事，敘述著角色、景物、事件、觀感、以及情緒等，藉由文字的組合詮釋以傳達各式各樣的主題。詞曲創作過程中，不論是「先有曲、後有詞」，或是「先有詞、後有曲」，詞曲風格經常相互影響(Wang, Syu, & Wongchaisuwat 2021)。由此可知，坊間多樣且豐富的流行歌包含著歌曲與歌詞，反映著人們的生活與文化；聆聽者不僅須使用耳朵聆聽樂曲節奏，往往也會透過文字解讀歌詞字裡行間的意涵。因此，藉由資訊科技解析歌曲和歌詞之間的關聯性，應該有助於音樂數位典藏(digital music archives)之知識庫建構(Raheb et al. 2022)。

音樂數位典藏是輔助教育場域的重要素材(Orio et al. 2009; De Valk et al. 2017; Raheb et al. 2022)，大多數對於音樂素材的歸類與分析著重於聽覺效果，而輕忽了音樂的文化意涵。在音樂風格分類方面，機器學習的模型已被廣泛應用(Bahuleyan 2018; Poonia, Verma, & Malik 2022)，其中卷積神經網路(Convolutional Neural Networks, CNN)常被用於構建分類模型(Aguiar, Costa, & Silla 2018; Nirmal & Mohan 2020; 郝沛毅等人 2020)。歌詞是音樂的文字表達，歌詞在音樂中可以深化聽眾對歌曲的理解和共鳴，透過歌詞分析方法可以深入研究和解釋歌詞的獨特意涵，常見的歌詞分析方法(Fang et al. 2017; Alexopoulos & Taylor 2020)如歌詞

主題分析(Wang et al. 2021)、歌詞情感分析(Hu, Downie, & Ehmann 2009; Jamdar et al. 2015)、文化和社會詮釋、以及敘事結構分析等(Smith 1980; Van Sickel 2005)。因此，在歌詞分析方面，為了解析歌詞的主題，可透過文字探勘 (text mining) 技術分析歌詞中的詞彙頻率、詞性、以及詞彙之間的關聯性等(Wang et al. 2021)。CombinedTM 模型能夠捕捉到具有連貫性的主題，並透過上下文學習使主題內容更加連貫(Bianchi, Terragni, & Hovy 2020)。綜合上述，本研究運用 CNN 模型進行歌曲風格分類，並採用 BERT (Bidirectional Encoder Representations from Transformers)嵌入模型(Devlin, Chang, Lee, & Toutanova, 2019)，結合 CombinedTM 主題模型(Bianchi, Terragni, & Hovy 2020)同時進行歌曲風格分類、以及歌詞主題探勘。本研究將探討以下三個問題：(1)如何運用 CNN 模型進行歌曲風格的分類；(2)如何使用 CombinedTM 主題模型，分析歌詞主題傾向；(3)根據歌曲風格分類結果與歌詞主題傾向分析，探討歌曲和歌詞之間的關聯性。

研究結果顯示，歌曲風格分類部分，本研究所選用之 CNN 模型能夠有效區別 Blues, Classical, Country, Disco, Hip-Hop, Metal, Pop, Reggae 以及 Rock 等九種歌曲風格，分類準確率(precision)高達 0.89，並能夠解析每一首歌曲之歌曲風格成分比例。歌詞主題部分，本研究所建構之主題模型結合資訊視覺化圖形(黃芝璇、馬麗菁 2021)可針對選用歌詞之文字意涵解析特定歌詞在「創意與音樂表現」、「社交和娛樂活動」、「日常和生活情境」、「情感和內心世界」、以及「感官和身體體驗」等五大主題的傾向。進一步，透過「歌曲類別與歌詞主題傾向雷達圖」分析，則可觀察出歌曲和歌詞之間的關聯性。

貳、文獻探討

一、音樂資訊檢索與 CNN 網路

音樂資訊檢索(music information retrieval, MIR)領域是一種多媒體檢索方式，通常針對音頻資料進行檢索(Yu et al. 2020; Elbir 2020; Singh & Biswas 2022)。常見的檢索方式包括根據作者資訊、音樂形式、創作組織或樂曲特徵(如調性、節奏)等進行檢索。音樂資訊檢索過程中，取樣(sampling)音頻資料時，通常不會將一長段音樂內容直接分析，而是將樂曲以不重疊的切片，取出較短音樂段，去除音樂片段中不具代表性的部分，促使模型處理具較高紋理特徵的圖像(Aguiar et al. 2018; Costa, Oliveira, Koerich, & Gouyon, 2013; Singh & Biswas 2022)。進一步，採用深度學習模型運用於歌曲風格分類，通常會先將音頻轉換為頻譜圖，再以 CNN 網路的深度學習模型(Lecun et al. 1998; Yu et al. 2020; Elbir 2020)進行音頻風格分類。此方法不需要人工定義特徵，可藉由捲積層、池化層、扁平層和全連接層的模型架構，將擷取到的特徵進行分類(Bahuleyan 2018）。

Bahuleyan (2018)指出使用 CNN 模型進行歌曲風格辨識之效能和準確率(accuracy)，均優於傳統機器學習演算法。然而，當參數增加、或卷積核的 VGG 網路架構更為複雜時，將會增加訓練時間成本。因此，本研究將採用參數較少、

捲積層深度較淺的 CNN 網路進行模型訓練，對音頻資料進行頻譜圖轉換並進行機器視覺分類。Yu et al. (2020)認為傳統卷 CNN 模型僅採用音頻頻譜圖分析，進行音樂風格分類，有可能違背聽眾對音樂聆聽時的心理感受。因而提出一種基於雙向循環神經網絡(attention mechanism based on Bidirectional Recurrent Neural Network, BRNN) 的結合注意力機制的模型(attention mechanism based on Bidirectional Recurrent Neural Network)，在相對多層的網路架構(五層以上)，能夠在音樂風格分類上取得較好的效果。

歌詞在眾多音樂作品中扮演重要的角色，在語義上具有重要性，Fang et al.(2017) 認為在音樂風格識別上整合整首歌詞的文字特徵能夠更精確的進行音樂風格分類，明確表示基於歌詞特徵對音樂資訊檢索的重要性。為了具有更為優異的音樂資訊檢索效果，Li et al. (2023)提出一種結合歌曲和歌詞資訊用於音樂風格分類的框架，在歌曲方面其使用 CNN 提取音頻特徵，在歌詞方面則使用了 BERT (Bidirectional Encoder Representations from Transformers)模型來獲取歌詞的語義信息。該研究結果顯示融合歌曲和歌詞的機制相較於僅採用歌曲或僅採用歌詞，在分類表現之準確率更為優異。

二、音頻與梅爾頻譜轉換

由於人耳的聽覺掩蔽效應(Auditory Masking)而導致人耳對於音頻的感知是非線性的，對於低頻感受強度會高於相同能量的高頻頻率(Wegel & Lane 1924)，隨著頻率增加，感受到相同音高的聲音會變得較弱。梅爾頻譜（Mel spectrogram）的轉換過程是一種基於此特性的非線性頻率轉換機制(Stevens, Volkmann, & Newman 1937)，透過快速傅立葉變換（Fast Fourier Transform，FFT），取得頻率域。例如，為了進行時頻分析(time–frequency analysis)，透過短時傅立葉轉換(short-Time Fourier Transform, STFT)將長時間音頻資料切分成許多等長、較短之音頻資料，再分別對這些音頻資料進行傅立葉轉換，用來描繪音頻資料在頻域、以及時域的變化(Singh & Biswas 2022)。經過 STFT 轉換的頻譜圖與三角帶通濾波器(Triangular Bandpass Filters)進行內積運算處理後所產生的梅爾頻譜圖之矩陣資料，可藉由取對數轉換為分貝單位(dB)，再轉換為類似人耳聽覺的梅爾刻度(Mel Scale)；隨之，運用可視化處理，繪製成梅爾頻譜圖。

進一步，可結合 CNN 進行影像識別，分析音頻特性。此轉換方式經常運用於語音辨識、語音合成、以及音樂信息檢索等領域(O'Shaughnessy 1987; Tzanetakis & Cook 2002)。

三、CNN 神經網路與音樂風格分類

CNN 神經網絡在機器視覺領域中表現出色，並在圖像識別和多類別分類等問題上取得了優異的成果(Yu et al. 2020; Elbir 2020; Singh & Biswas 2022; Li et al. 2023)。亦有研究指出運用梅爾頻譜圖結合 CNN 架構(Bahuleyan 2018)，相較於傳統機器學習方法(如邏輯回歸、隨機森林、以及支持向量機等)，具有相對優異的

表現。Lecun et al.（1998）指出 CNN 的捲積核具有平移、翻轉、縮放和光照不變性，使得模型能夠對不同環境和背景下的影像進行準確識別和學習。相比之下，對於沒有卷積核的全連接神經網絡（FNN）而言，圖像內容的變化將導致完全不同的輸出，需要重新學習，這使得無法捕捉特徵，效率低下，甚至無法學習。隨著 CNN 圖像識別在機器視覺領域中的技術發展和改進，已衍生了許多變異模型，如 VGG（Simonyan & Zisserman 2014），LeNet（Lecun et al.1998）和 GoogleLeNet（Meyers et al. 2015）等。

由於音樂作品的風格分類種類繁多，因此模型設計過程可先經由扁平層轉換多維度張量（tensor）輸出（output）成為一維度張量後，再進行分類處理。處理此類問題較常使用 SoftMax 激活函數作為輸出，並結合採用 Categorical Cross-Entropy（CEE）損失函數 (Aguiar et al. 2018; Doon, Rawat, & Gautam 2018; Lecun et al. 1998)。SoftMax 激活函數可針對某一類別，加總所有可能的類別特徵的機率，其機率總和為 1。以機率方式來統計神經網路的輸出結果，最大值即代表預測結果。SoftMax 函數表示方式如式 1：

$$y_i = \frac{e^{z_i}}{\sum_{j=0}^{C} e^{z_j}} \qquad \text{式 1}$$

y_i代表為預測機率，C 代表為類別總數，z 為神經網路的預測值，i 為神經網路第 i 個輸出，j 與 i 意義相近，同為索引值，表示為神經網路第 j 個輸出。

多類別分類問題經常採用 Categorical Cross-Entropy（CEE）損失函數，其使用 SoftMax 函數，並結合 Cross-Entropy 損失函數。因此，可計算每個類別的預測機率、以及與真實值之間的誤差。式 2 表達 CCE 函數計算方式，其中，y 為預期輸出結果，$\widehat{y_{i,j}}$為 SoftMax 函數輸出結果，f為 SoftMax 函數，C 代表為類別總數，N為批次數量。

$$CCE = -\frac{\sum_{i=1}^{N}\sum_{j=0}^{C} y_{i,j}\log\left(f\left(\widehat{y_{i,j}}\right)\right)}{N} \qquad \text{式 2}$$

四、歌詞文本與主題模型

歌詞可以提高音樂分類系統的分類準確性，Hu et al. (2009)運用歌詞、音頻、以及結合歌詞和音頻等三種不同特徵選擇方法，構建音樂情緒分類機制，研究結果發現音頻特徵不一定會優於歌詞特徵，整合歌詞和音頻特徵可以改善音樂情緒分類機制的性能。Wang et al. (2021)提出了一種音樂自動標籤機制，採用卷積神經網絡(CNN)和循環神經網絡(RNN)對於歌詞文字及結構特徵加以提取後，並結合音頻特徵加以訓練之分類器，其在性能上優於先前研究中僅使用音頻的分類方法。Fell et al. (2023)採用包含歌詞與歌曲的大型歌曲語料庫(WASABI)，設計一自動歌詞標註機制可用於處理歌詞的結構分段、確立歌詞主題、歌詞摘要、以及

歌詞情感分析等，還可進一步將分析結果轉化為音樂知識圖。透過音樂知識圖的分類和歌曲推薦功能，使用者將可根據使用情境的需求，搜尋、瀏覽大量的歌詞與歌曲。由上述研究可知，實現有效的音樂檢索功能，歌詞與歌曲的特徵擷取以及應用場景至關重要。

BERT 是一種預訓練的語言模型，能夠生成單詞和句子的表示形式(Devlin et al. 2019)。Sentence-BERT（Reimers & Gurevych 2019）是在 BERT 的基礎上進行改進的模型，廣泛應用於自然語言處理（NLP）領域，用於解析句子的語義，常用於文本分類、句子相似度、文章分群和資訊檢索等任務。Sentence-BERT 主要用於句子級別的表示，旨在掌握整個句子的含義。該模型通常使用兩種訓練方法：孿生網路（Siamese network）(Bromley, Guyon, LeCun, Säckinger, & Shah 1993)和三元組網絡（Triplet network）(Hoffer & Ailon 2015)。孿生網路由兩個共享權重的子網絡構成，透過計算最小損失（loss）來評估兩個輸入之間的相似度。模型接受兩個句子作為輸入，並學習預測這兩個句子在語義上的相似性，該方法常用於句子相似度計算和句子分類。三元組網絡的訓練樣本包含三個部分：錨定句子（anchor）、正例句子（positive，表示與錨定句子語義相似）和負例句子（negative，表示與錨定句子語義不相似）。透過學習將正例句子與錨定句子靠近，將負例句子與錨定句子遠離，該模型利用損失函數最小化正例句子與錨定句子之間的距離，同時最大化負例句子與錨定句子之間的距離，從而最小化三元組損失（triplet loss），以實現詞句檢索或分群等任務。

ProdLDA 是一種基於 LDA（Latent Dirichlet Allocation）所發展出來的主題建模的模型(Srivastava & Sutton 2017)，採取機率模型將文件集合中的每個文件表示為多個主題的混合體，藉此建構文件與主題之間的關係，以強化主題建模的性能。ProdLDA 的訓練過程與 LDA 類似，透過對文件集合進行迭代採樣，可以得到每個文件與主題之間的分佈、以及主題和詞之間的分佈(Kingma & Welling 2013)。進一步，相較於傳統的 LDA 模型，ProdLDA 考慮了主題之間的時序關係，引入了一個時間變量，用於表示主題的演化過程，透過生產過程和時間變量，更能夠更準確地解析文本中的主題結構和主題演化，從而提供更有價值的主題模型。在應用上，ProdLDA 不須複雜程序，毋須考慮模型內部邏輯，即可為特定文本素材進行模型設計。因此，可以更完善地解析文本中主題的變化趨勢，以運用於歌詞文本分析、文本摘要、以及主題傾向檢測等任務。

Bianchi, Terragni, & Hovy (2020)結合了 ProdLDA (Srivastava & Sutton 2017)以及 Sentence-BERT (Reimers & Gurevych 2019)的特點，建構出一組合式主題模型 Combined Topic Model (CombinedTM)。該組合式主題模型的運作，先使用 Sentence-BERT 預訓練模型來快速取得句子的嵌入特徵，由此可使用機率的方式提升自然語言中語法的正確性，並透過詞向量的方式捕捉詞與詞之間的脈絡關係，用來生成連貫且有意義的文本。然後，將文本取得嵌入投影至隱藏層後並與口袋字特徵相連接，再使用 ProdLDA 將處理後的文本轉換為潛在特徵（latent representation）並使用解碼器網路(Neural Variational Document Model,

NVDM)(Miao, Yu, & Blunsom 2016)從潛在特徵中生成口袋字，並運用專家乘積值進行單詞與主題的整合，可應用歌詞文本的主題提取。

參、研究方法

一、歌曲風格分類模型設計

首先，必須取得歌曲音頻資料進行訓練。Tzanetakis & Cook (2002)收集當時坊間多樣的音樂素材，藉由音色紋理(timbral texture)、節奏內容(rhythmic content)、以及音高內容(pitch content)等特徵，來訓練音樂風格分類器。為了確保不同的錄音品質差異的兼容性，其所收集的音樂素材有來自於收音機、CD、以及 MP3 等音頻，並採以 22050 Hz、16 位元(bits)深度、單聲道的格式存儲，並且分別使用 100 個代表性的片段(每片段 30 秒)進行訓練。此 GTZAN 音樂素材集中，將音樂風格分成十種類別，包括：Blues, Classical, Country, Disco, Hip-Hop, Jazz, Rock, Reggae, Pop 以及 Metal 等；音樂素材內容的構成極為多元，包含純樂器演奏、以及人聲與樂器演奏的融合等形式(Tzanetakis & Cook 2002)。由於此 GTZAN 音樂素材集具備上述的多樣性與真實性，所以本研究使用 Tzanetakis & Cook (2002)提供的音頻素材為歌曲風格的訓練資料。然而，GTZAN 部分音頻素材具有標籤誤植的情形，而且於 Jazz 風格類別的第 54 個樣本檔案也有壞軌的狀況(Sturm 2012)。因此，本研究捨去該資料集的 Jazz 類別，保留其餘九種風格進行分類訓練。為了增強圖像識別紋理特徵(Costa et al. 2013)，音頻資料前處理時將 GTZAN 資料集樣本分割為 3 秒一個片段。由於每首歌長度均等長為 30 秒，所以每一首歌會被分割成 10 個片段；然而，每一種音樂風格都包含 100 首歌，因此經過分割後每一種風格的音樂擁有 1000 個切片樣本，前處理後總共 9000 個音頻資料樣本。

針對歌曲風格分類機制本研究採用 CNN 神經網路進行歌曲風格的訓練學習與分類，建構歌曲風格分類模型流程如圖 1 所示。首先，從 GTZAN 取得音頻資料並將音頻資料切割成每三秒鐘為一個片段；隨即，進行這些音頻切割片段的前處理，包含資料增強(data augmentation)、將增強後歌曲片段轉換為梅爾頻譜、以及將梅爾頻譜圖進行標準化處理。訓練階段，先將所有歌曲片段依 9:1:1 分成訓練集、測試集、以及驗證集，並使用超參數（hyperparameter）方式進行模型訓練，然後進行效能評估。最後，使用訓練完成模型對未標記資料進行預測。資料增強方法可用來改善小樣本數訓練資料集的限制(Litjens et al. 2017)，透過資料增強使得樣本更加多樣化，資料增強的手法包括影像縮放、添加雜訊、旋轉影像、顏色調整、光源調整等(Shorten & Khoshgoftaar 2019)。然而，資料增強方式上有些限制；舉例來說，當處理手寫數字辨識資料集，若將數字進行旋轉方式資料增強，可能會導致特徵與其他的數字發生混淆(Lecun et al. 1998)。

模型架構採用 4 個捲積層，皆使用 ReLU 激活函數（activation function），且分別都連接 2×2 的池化層。全連接層輸出藉由 softmax 激活函數，將輸出結果轉為機率形式，損失函數使用 Catagorical Cross-Entropy，總迭代次數為 40 代。最

後，透過扁平層將資料壓縮成一維特徵張量，並將其傳入全連接層進行歌曲風格分類。

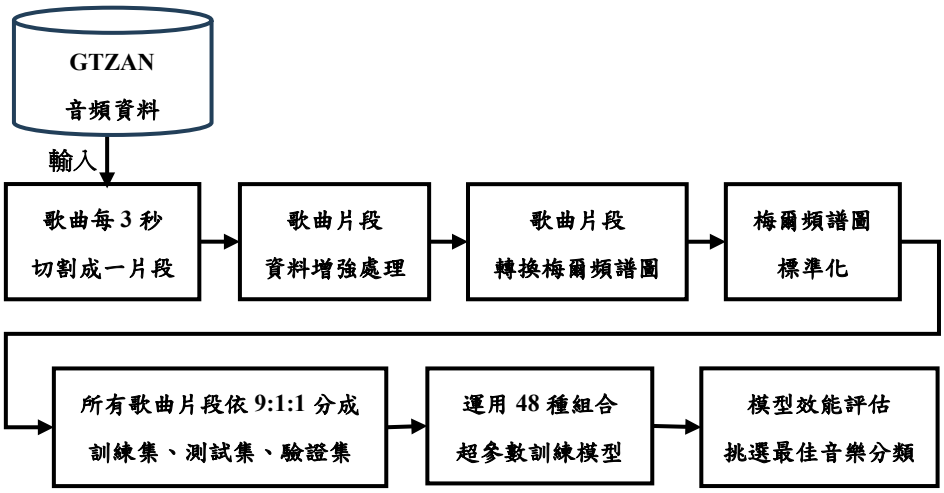


圖 1：建構歌曲風格分類模型之流程

二、歌曲音頻處理與梅爾頻譜圖轉換

根據上述流程，實驗後發現，未加入增強資料的模型會產生過擬合現象（overfitting），使得訓練資料的分類表現較好，而無標記測試資料上的分類表現較差。主要的可能原因包含：(1)訓練資料量不足；(2)參數設置過於複雜；(3)訓練資料多樣性不足等原因。然而，過擬合將造成模型泛化能力不足，因而降低無標記資料分類效果。

本研究的歌曲音頻資料屬於時間序列，其音頻先後順序尤其重要，若時間先後關係發生改變則會造成破壞。因此，本研究在資料增強策略上包含：(1)隨機加入不同強度的白噪音(0～0.5 之間的亂數為參數值)。(2)利用時間伸縮，將音頻資料播放速率進行隨機調節(以調慢一倍到加速一倍之間隨機速率選擇，使得所產生的圖形比原始資料更長或更短)。(3)國際慣用音律為十二平均律，即將一個八度音間平均切分為 12 個半音階，當第 12 個半音時為高八度，聽覺上會與原曲調一樣，但頻率上會變為原本的 2 倍；因此，採用 0～11 之間的整數亂數隨機選擇樂曲調性於 12 個不同音律。(4) 隨機音頻相位極性反轉，使聽覺感受失去立體感；採用 0 或是-1 隨機整數亂數，當選擇-1 則將聲波相位反轉。(5)利用隨機增益效果，以 0～5 之間的隨機浮點數，對音頻資料進行隨機增益調整，創造多樣的音量變化。資料增強策略的執行方式，採用 python 套件 librosa (McFee et al. 2015)，每次隨機選擇一種方法，對原始音頻資料進行處理後，再轉換成梅爾頻譜圖以進行模型訓練。實驗資料透過資料增強後的梅爾頻譜圖，如圖 2 所示，其中圖 2(a)為未進行資料增強策略的原始梅爾頻譜圖(original signal)、圖 2(b)為加入白噪音的梅爾頻譜圖(add white noise)、圖 2(c)為利用時間伸縮隨機調節速率的梅爾頻譜圖(time stretch)、圖 2(d)為隨機選擇樂曲音律的梅爾頻譜圖(pitch scale)、圖

2(e)為隨機音頻相位極性反轉處理的梅爾頻譜圖(invert polarity)、以及圖 2(f)為隨機增益效果處理的梅爾頻譜圖(random gain)。

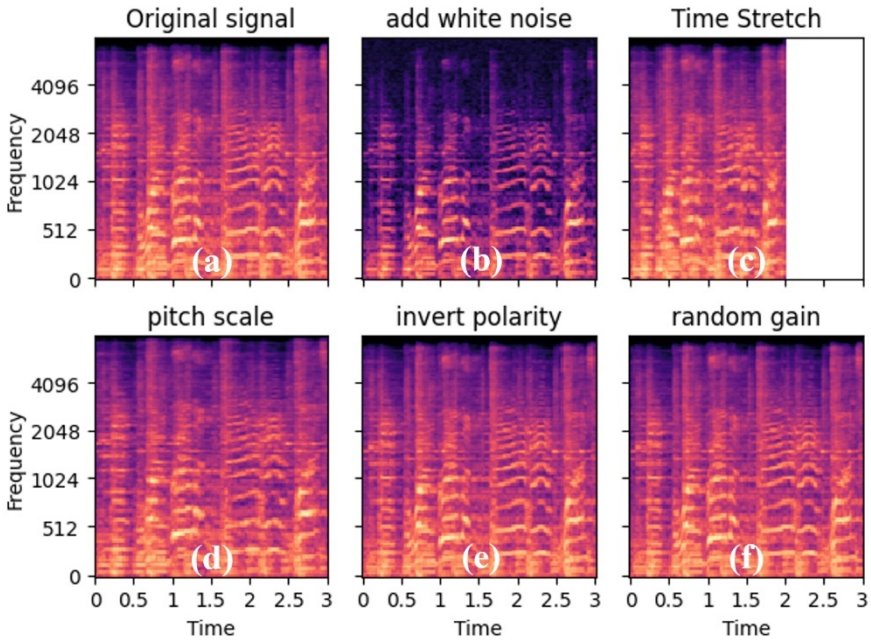


圖 2：音頻資料增強後之梅爾頻譜圖

三、歌曲風格分類模型效能評估

本研究之訓練集、測試集和驗證集的切割比例為 9:1:1。其中，訓練集包含 8100 張頻譜圖，測試集和驗證集各包含 450 張頻譜圖。這些頻譜圖的尺寸為 288(寬)×432(高)，並將圖像中的像素值除以 255 進行標準化處理，即當圖像輸入形狀(shape)參數為(288, 432, 4)，表示長度 288、寬度 432、深度 4。圖像的色彩模式為 RGBA (Red Green Blue Alpha, RGBA)，其中 Alpha 通道代表透明度參數。同時，訓練過程中使用的批次大小(batch size)為 128。為了比較不同超參數組合對模型效能的影響，本研究使用了以下不同的超參數組合進行實驗：(1)是否使用影像增強資料(使用影像增強/不使用影像增強)；(2)是否使用網路輸出層標準化演算法（Batch Normalization）(使用 Batch Normalization/不使用 Batch Normalization)；(3)使用 4 種不同的學習率（learning rate：0.00005、0.001、0.01、0.03）；(4)使用 3 種不同的權重初始化方法：RandomNormal = 0.01、glorot uniform (Glorot & Bengio 2010)、he normal (He, Zhang, Ren, & Sun 2015)，總共有 48 種(2×2×4×3)不同的模型訓練組合。

在超參數訓練過程中，若精確率(precision)超過總迭代次數的 1/3 後(即連續 12 次未出現提升)，則立即停止訓練。最後，藉由 Softmax 函數輸出的機率數值最大值可判定資料所屬的分類項目。模型的輸出結果共涵蓋 9 種歌曲類別，為了驗證模型泛化性以及強健性（robustness），本文參考具指標性的音樂串流平台 Apple Music 分類，其分類方式為每首作品給定單一類別標籤代表其歌曲風格。該平台之標籤分類分別為 Blues, Classical, Country, Disco, Hip-Hop, Metal, Pop,

Reggae 以及 Rock 等共九種類別，跨足多個年代(西元 1930~西元 2020)，其中 Classical 類別的音樂絕大部分為純器樂演奏之曲目並沒有歌詞。於是，九種歌曲類別每類各下載 10 首歌曲，合計彙總 90 首歌曲。由於本研究著重於探究歌曲與歌詞的特質，所以在歌詞分析部分捨棄不具備歌詞的項目後(如 Classical 音樂類別為純樂器演奏沒有歌詞、以及 Disco 類別中有兩首舞曲沒有歌詞)，合計共採用了 78 首歌詞。

實驗時，本研究共使用了上述 48 種組合來訓練模型，為了挑選最適合的模型，於是運用視覺化工具 Tensor Board (https://www.tensorflow.org/tensorboard)來觀察各模型的效能表現，部分訓練模型的準確率(accuracy)表現如圖 3 所示(橫軸表示 epoch，縱軸表示 Accuracy)。觀察發現，準確率較高的模型通常都經過了資料增強和 Batch Normalization 的處理(其中僅 30 號和 34 號模型未使用 Batch Normalization)。此外，使用的學習率參數較低可訓練出相對更準確的模型，然而也降低模型的收斂速度。訓練過程中，在大約第 25 到第 30 個迭代期間，模型的準確率並沒有持續往上提升。綜合觀察超參數調整對模型性能的影響，發現資料增強模式、Batch Normalization、以及學習率等因素將直接影響模型的性能；然而，權重初始化對模型性能的影響並不顯著。

為了避免所選用的模型因過度擬合的現象，導致雖具有高準確率，但泛化能力卻不佳，而導致對多樣化的歌曲進行音頻資料分類時的準確率明顯下降。因此評估模型效能時，應該選擇較小損失(loss)的模型(即最後一次迭代的 loss 值)，同時評估模型的泛化能力和準確率。因此，本研究先捨棄準確率低於 85%的模型，然後再挑擇了 45 號和 41 號模型作為候選模型(附表 A.1)。相較於 Tzanetakis & Cook (2002)對十種音樂類型的分類準確率表現(0.61)，以及 CNN VGG16 (卷積層層數為 16 層)網路架構(Bahuleyan 2018)之分類準確率表現(0.89)，本研究僅使用 4 個卷積層架構之 CNN 網路，透過資料增強策略結合超參數訓練，亦能使得歌曲風格分類精確率高達 0.89。

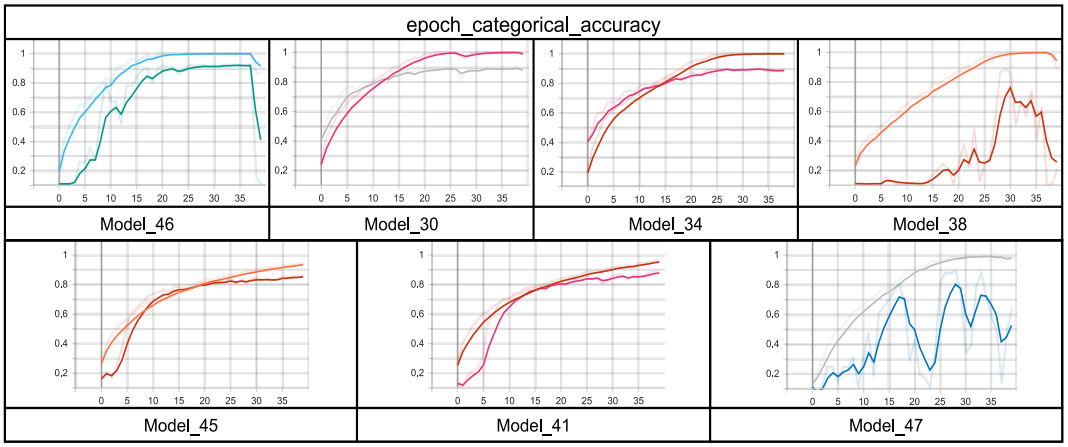


圖 3：不同參數組合訓練的模型效能

四、歌詞主題模型建構

本研究採用 CombinedTM 主題模型，對歌詞文本進行分析。首先，採用 Python 的自然語言處理工具 NLTK 套件，進行歌詞文本斷詞、斷句、標記化、詞幹提取、以及語意推理等處理（Loper & Bird 2002）。藉此擴增停用詞(stop words)，將歌詞中沒有意義的狀聲詞(例如 Ah, Whoo, yeah 等)加入停用詞。進一步，運用 CombinedTM 主題模型中的 Sentence-BERT 計算字詞向量，並使用 ProdLDA 模型擬合主題和口袋字向量，以生成一系列基於生成主題的口袋字。最後，將這些不同主題的口袋字輸入 ChatGPT，要求它分別為這些口袋字賦予最適當的主題名稱和情緒特徵。上述的歌詞主題定義步驟，如圖 4 所示。

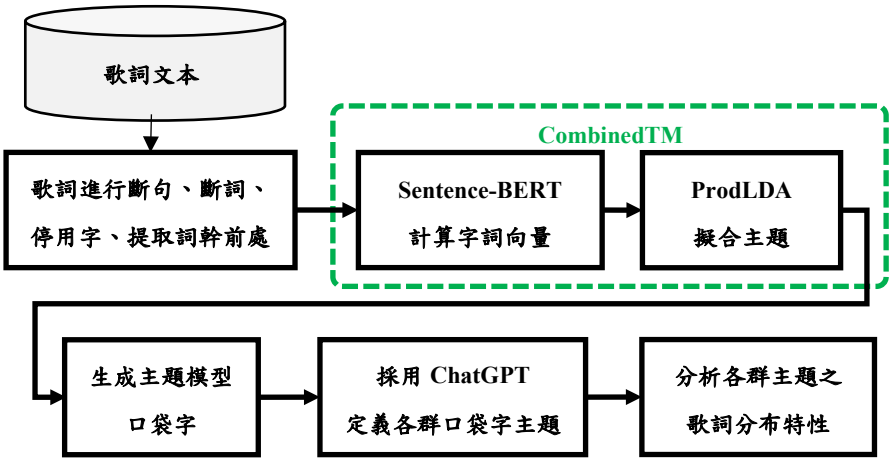


圖 4：主題模型與歌詞主題建構之流程

CombinedTM 主題模型無論主題數量設置多寡，大都具有一致性（Bianchi, Terragni, & Hovy 2020; Bianchi, Terragni, & Hovy2020; Qiang, Qian, Li, Yuan, & Wu 2020）。由於本研究所使用的歌詞樣本相對較小，僅包含 78 個文檔。在決定採用多少個主題數量部分，對於大多數使用主題模型的應用，選擇理想主題數量一直是此類分析中經常存在的問題(Zhao et al. 2015; Vayansky & Kumar 2020)。常見的作法是使用相異 k 個主題，運用迭代反覆測試資料，然後分析某些指標來比較結果而決定。然而，這樣的過程不僅會佔用大量的時間和計算成本，而且可能產生過於龐大的主題數量來進行歸納分析(Greene, O'Callaghan, & Cunningham 2014; Churchill & Singh 2022; Sbalchiero & Eder 2020)。因此，本研究藉由 CombinedTM 主題模型將其歸納為 5 個主題並採用雷達圖的概念，以視覺化方式呈現歌詞文本的主題的傾向。對於使用者而言，將能避免因過多主題資訊而造成資訊過載(information overload)，或是因為主題數量太少使之難以衡量歌詞主題與歌曲風格間的關聯(Hwang & Lin 1999)。於是採 CombinedTM 模型使用深度學習方法對主題內容進行特徵提取的訓練，並針對該模型的參數設置包括：bow_size = 1999（經過前處理後的詞彙數量）、contextual_size = 768（Sentence-BERT 嵌入向量的維度）、n_components = 5（指定五個主題數量）、num_epochs= 10（訓練迭代次數）、model_type = "prodLDA"（主題模型設置為 prodLDA）、

hidden_sizes = (100, 100)（隱藏層大小為 100）、activation= "softplus"（啟動函數設置為"softplus"）、dropout = 0.2、lr = 2e-3（學習率設置為 0.002）、以及 solver = "adam"（優化器設置為"adam"）等。透過上述 CombinedTM 模型的參數設定，最終生成了五個主題，每個主題包含了不同數量的關鍵詞，可以將這些詞視為各主題的代表性概要。

最後，針對每個主題從中選擇了最具代表性的 30 個候選詞作為主題的摘要，並使用 OpenAI 的 GPT-3.5 語言模型 ChatGPT（Brown et al.2020）為這五個主題的候選詞（即主題的摘要）進行命名(Gao et al. 2022; 2023)。根據 ChatGPT 提示語(附表 A.2)所對應之主題名稱，整理如表 1 所示。

表 1：歌詞主題傾向指標與候選關鍵字

主題指標	**候選關鍵字**（前 30 候選詞）
創意與音樂表現	fun, metaphor, lord, turned, rollin, spit, singing, time, hard, else, tickin, fraud, character, chasin, get, hundred, bit, chandelier, brutal, run, ring, raining, little, song, meal, life, care, trapped, chirp, cartier
社交和娛樂活動	blow, wide, squeeze, want, another, whatever, sometime, sing, gore, leg, someone, pushed, funkytown, patience, mercy, trick, kissed, class, teeth, stabbin, taking, bad, around, bought, tuesday, caught, wide, dipper, hittin, enjoy
日常和生活情境	hound, name, bad, country, steamship, case, house, start, man, work, dog, sex, brag, high, could, alright, virginia, fleshy, gather, ragamuffin, quit, tender, sweet, bawl, month, rate, left, fare, drunk, brave
情感和內心世界	look, schizo, born, die, chandelier, glitter, ring, coolie, knew, feeling, latest, arizona, tonight, night, kid, miss, smiling, whip, sword, badge, bawl, happy, well, feel, levitating, continuous, lazy, said, anything, strength
感官和身體體驗	death, work, box, bedroom, skintight, longing, candle, trigger, bigger, maggie, heart, rollin, understand, sea, ticket, stick, heat, basement, uptown, hound, warm, trapping, cream, quaking, panic, movin, see, smile, steal, fallin

肆、實驗評估與探討

一、候選模型評估與歌曲風格分析

為了評估候選模型第 41 號及第 45 號之歌曲風格分類效果(附表 A.1)，將測試集輸入模型並採用混淆矩陣(confusion matrix)進行成效評估(附圖 A.1)。第 41 號模型對測試集的分類表現呈現於圖 5，其準確率(accuracy)為 0.87，發現其標示為 Country 歌曲風格的少數歌曲片段被歸類到 Pop, Rock, Disco, Reggae 等風格，Country 類別之分類準確率為九個類別中相對較差(0.82)。除了 Rock(0.72)之外，其它各類別的分類表現之準確率則大都高於 0.82 (0.84~0.98)，例如 Pop 的準確率為 0.84、Reggae 為 0.87、Classical 達 0.98、以及 Metal 為 0.89，表示第 41 號模

型的歌曲風格分類有不錯的表現。圖 6 則呈現第 45 號模型的分類表現，該模型的準確率為 0.85，其分類錯誤部分相較於第 41 號模型更為分散；例如，其中表現最差的分別為 Country, Rock, Hip-hop，其準確率分別是 0.72 以及 0.73，其它歌曲風格分類的準確率大都超過 0.85 以上，甚至到 0.98，整體表現亦不差。然而，仔細觀察這兩組模型都具有一個共通特性，具備 Country 歌曲風格的些許歌曲片段會被納入到 Pop, Rock, Disco, Reggae 等，而且被納入 Rock 比重都是最高。換句話說，Country 與 Rock 這兩種歌曲風格，極有可能因為彼此音樂風格相互影響而相似，進而造成準確率偏低。

歸納其原因，在 1960 年至 1970 年代，美國西部的鄉村搖滾音樂文化興起，這導致了一系列的變革(Pruitt 2019; Martinez 2021)。這個時期的音樂跨界結合打破了種族隔閡，也為非主流音樂進入音樂市場提供機會。當時音樂創作者自由創作歌曲，經常將生活、偏好、文化、民族、以及歷史等元素融入歌曲(賴靈恩 2015；鄭英傑 2022)。許多鄉村音樂歌手嘗試將鄉村音樂、搖滾樂、以及流行音樂等風格結合，創造出融合了搖滾音樂和鄉村音樂的曲風。例如，加拿大創作型歌手艾琳·雷吉娜·仙妮亞·唐恩（Eilleen Regina Shania Twain）曾在鄉村音樂和流行音樂領域取得成就，後來更融合鄉村、搖滾、和流行音樂等元素，使得鄉村音樂在 90 年代進入了主流音樂市場(宋英維 2011)。

為了驗證上述觀點，本研究根據 Apple Music 串流平台所定義的 Country 歌曲風格類別，隨機取得十首歌曲，每首歌會被分割成 3 秒鐘一個片段，透過第 41 號模型對 Country 風格之歌曲進行分類評估時，可以觀察出 Country 風格歌曲實際上有部分比例之歌曲片段會被分配到其它歌曲風格類別中。如圖 7 所示，橫軸為被分配之歌曲風格類別，縱軸表示該歌曲片段所對應類別之占有數量。其中，有將近 40%的歌曲片段會被歸類於 Country；然而，亦有近 32%的歌曲片段會被歸類於 Rock；10%的歌曲片段會被歸類於 Matel，近 8%的歌曲片段會被歸類於 Classical, Pop 或 Reggae；甚至，還有少數歌曲片段會被分類於 Hip-Hop 等類別。綜合圖 7 與圖 8 呈現的歌曲風格之音樂成份分析，得知音樂藝術著重整體感，當某些風格片段出現的比重明顯突出於其它風格時，便會影響聽眾對該歌曲風格的認知。並驗證了，該時期的鄉村音樂與搖滾音樂的確相互影響，這樣的文化對後來的音樂發展亦產生了深遠影響，不僅推動了不同音樂風格的融合，也開啟了對多元文化和多樣性的重視(Martinez 2021)。

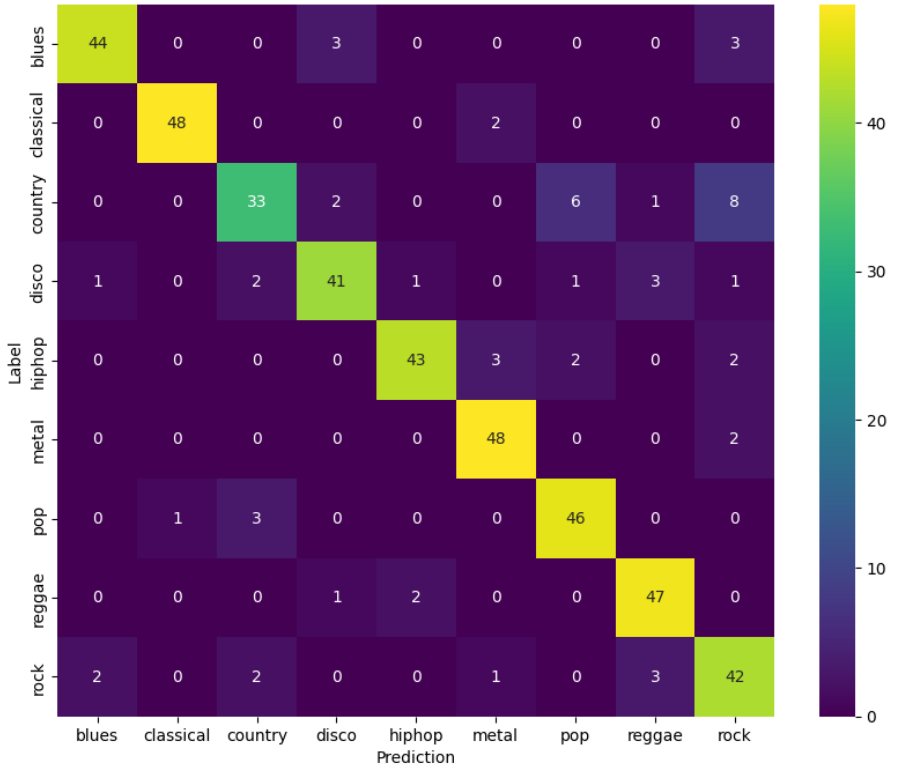


圖 5：第 41 號模型的混淆矩陣

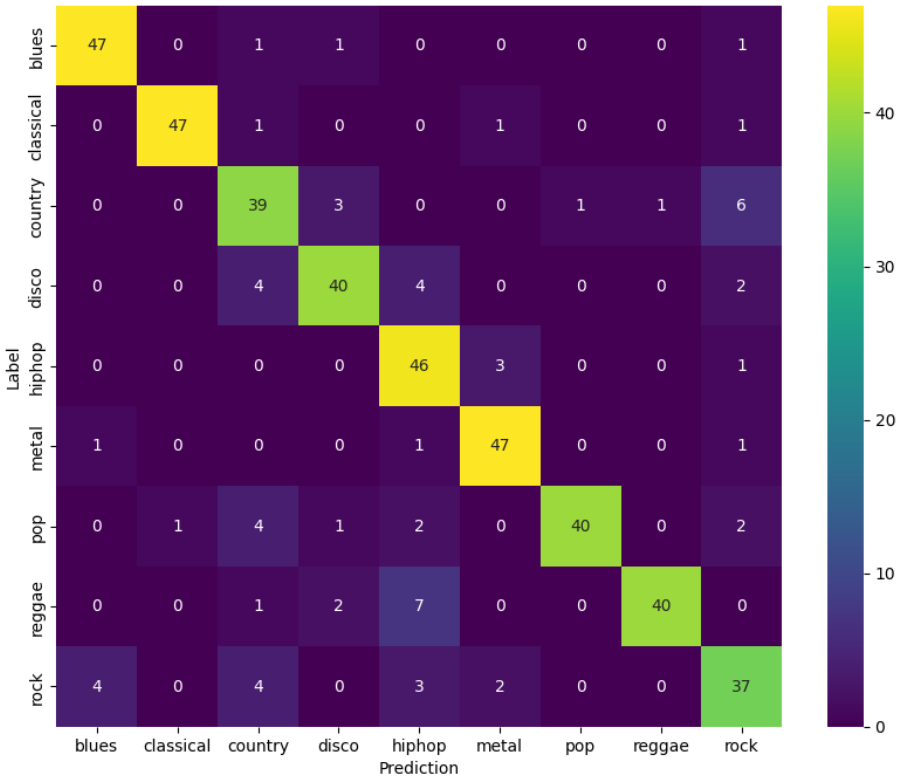


圖 6：第 45 號模型的混淆矩陣

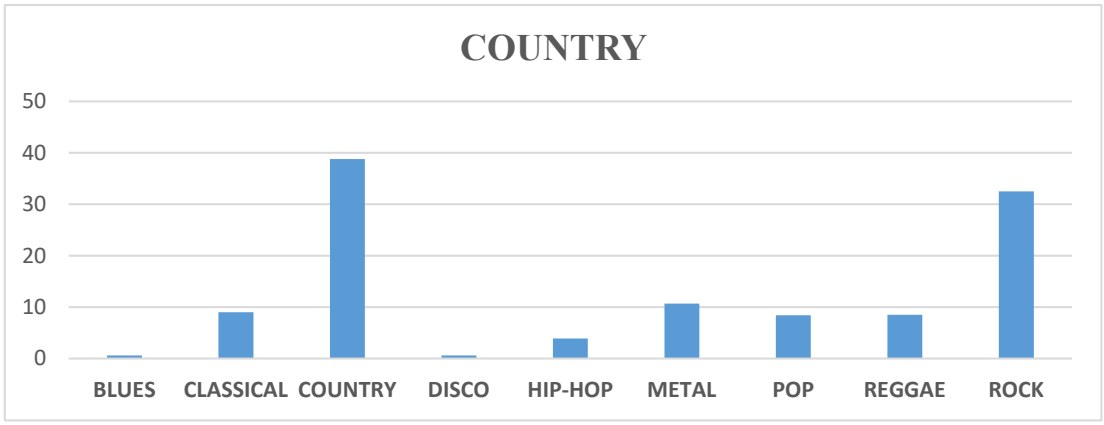


圖 7：採用第 41 號模型評估 Country 歌曲風格之音樂成份

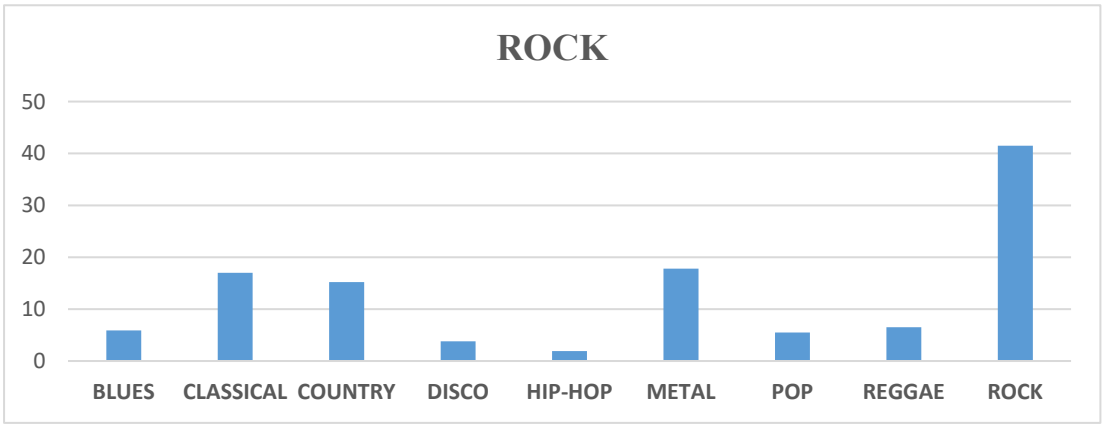


圖 8：採用第 41 號模型評估 Rock 歌曲風格之音樂成份

二、歌詞主題傾向分析

為了試圖了解歌曲風格與歌詞之間的脈絡關係，本研究進一步根據前文所定義的歌詞主題傾向指標與候選關鍵字(表 1)，將不同歌曲風格所對應的歌詞進行歌詞主題傾向分析。由於該些英文歌詞為了配合曲子的律動，經常會配合英文押韻或旋律而調整句子結構，抑或是在歌詞中加入大量的狀聲詞與語助詞，例如 Ah, Woooo, yeah 等英文字；為了避免這些字詞對於歌詞的主題傾向分析產生不當影響，本研究對每一首歌的歌詞皆進行文字前處理(preprocessing)，包含中將這些歌詞進行斷句、斷詞、以及去除對主題不重要的停用字(stop word)。歌詞主題傾向分析乃基於歌詞文本進行字頻(term frequency)計算，參照表 1 的歌詞主題傾向指標與候選關鍵字之對應關係；藉此觀察每一首歌詞的字詞有多少比率分別會對應於五個不同歌詞主題指標的候選關鍵字，以機率形式呈現每一首歌的歌詞之主題傾向。換句話說，每一首歌詞等同一篇文章，不同的文章特性應會著重於不同的議題；藉由五個主題指標(表 1)，將可用來衡量歌詞的文本主題傾向。

本研究彙整前述 Apple Music 串流平台所蒐集之 Blues, Country, Disco, Hip-pop, Metal, Pop, Reggae, Rock 等八種歌曲風格的歌詞(因 Classical 類別為古典音樂並無歌詞，因此不納入歌詞分析)，經文字前處理後，進行歌曲類別之歌詞主題傾向分析。如圖 9 所示，該些作品之歌詞彙整後的整體歌詞(Overall)主題傾向分

別可透過這五個指標平均值進行評估，其中「創意與音樂表現」的得分為 0.13、「社交和娛樂活動」的得分為 0.23、「日常和生活情境」的得分為 0.19、「情感和內心世界」的得分為 0.31、以及「感官和身體體驗」的得分為 0.15。也就是說，根據所取樣的歌詞內容，大多數的創作著重於「情感和內心世界」，其次為「社交和娛樂活動」，再來才是「日常和生活情境」等主題。為了便於觀測與判斷每一種歌曲類別對應於歌詞主題傾向的狀態，透過標準化 Z 分數(Z score)來描述八種歌曲類別對應整體歌詞(Overall)主題傾向差異量，以更全面地呈現結果，如附表 A.4 所示。

Overall	創意	社交	日常	情感	感官
Blues Mean	0.15	0.18	0.21	0.34	0.12
Country Mean	0.17	0.20	0.22	0.28	0.14
Disco Mean	0.12	0.17	0.18	0.41	0.12
Hip-pop Mean	0.12	0.24	0.22	0.29	0.14
Metal Mean	0.09	0.29	0.17	0.20	0.25
Pop Mean	0.14	0.21	0.20	0.32	0.12
Reggae Mean	0.14	0.36	0.16	0.24	0.11
Rock Mean	0.11	0.18	0.17	0.36	0.18
Mean	0.13	0.23	0.19	0.31	0.15
Min	0.09	0.17	0.16	0.20	0.11
Max.	0.17	0.36	0.22	0.41	0.25
Std.	0.02	0.06	0.02	0.06	0.04

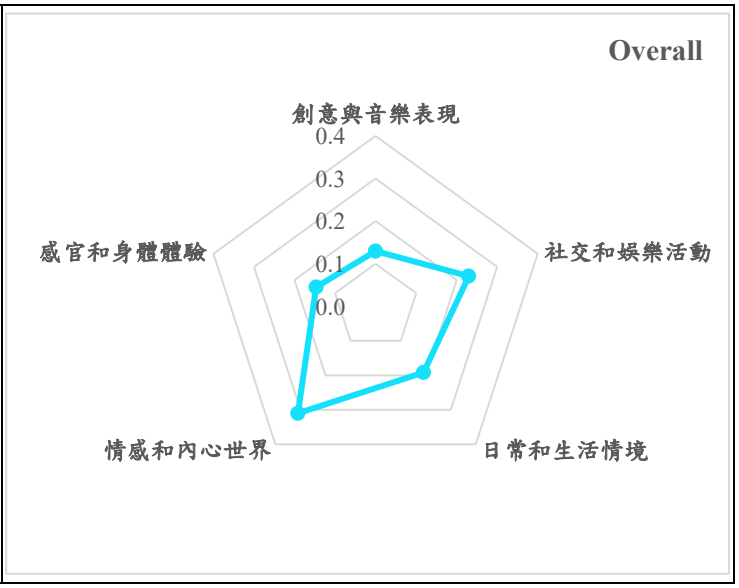


圖 9：彙整所有歌曲類別(Overall)之歌詞主題傾向雷達圖分析

進一步探究圖 10 的 Blues 之歌詞主題傾向分析，其中「創意與音樂表現」的得分為 0.15、「社交和娛樂活動」的得分為 0.18、「日常和生活情境」的得分為 0.21、「情感和內心世界」的得分為 0.34、以及「感官和身體體驗」的得分為 0.12。以 Blues 單獨類別的歌詞相較於整體歌詞(Overall)，「情感和內心世界」的主題成分占比相對更高，其值為 0.34，文本充斥著憂鬱與憤怒的語詞。此現象可追溯至藍調音樂的起源，藍調(Blues)音樂源於 19 世紀初非洲裔美國人的文化背景，這些黑人被迫至美洲大陸拓荒或從事勞務，經常受到雇主剝削，且人權自由意識逐漸抬頭；因此，這些黑人創作者利用工作閒暇之餘歌唱抒發身心壓力，排解受到迫害的心靈狀態，所以歌詞中的字句經常夾帶些哀怨情緒以反映當下的社會和政治問題所帶給他們的苦難(李怡萱, 2019)。藉由「情感和內心世界」指標所對應之候選關鍵字(表 1)，可觀察出此主題中出現了有關 look(注視)、schizo(精神分裂)、born(出生)、die(死亡)、coolie(苦力)、feeling(感覺)、night(夜晚)、kid(孩子)、miss(思念)、smiling(微笑)、whip(鞭子)、sword(劍)、badge(徽章)、 bawl(放聲痛哭)、levitating(懸浮)、continuous(連續的)、lazy(懶惰的)、anything(任何事情)、以及 strength(力量)等字詞。這些字詞充滿了對奴隸文化、困境、種族歧視、暴力、以及社會不公等殘酷現實的隱喻，同時也傳達了面對生活中的困境和挑戰，嘗試在成長過程中改變自己的處境，並為公平和正義而奮鬥，尋找出路的渴望(Narváez

1994; Stewart 2005; Byrd 2023)。列舉於 Blues 類別第 5 首，由羅伯特·強森（Robert Johnson）於 1936 年創作的《Cross Road Blues》的歌詞，整首歌描繪主人翁在十字路口的遭遇，抒發了對困境、絕望、無助、以及孤獨的情感，也反映了創作者所經歷的挑戰和不公(Compagna 2001; Rothenbuhler 2007)。

在圖 10 的 Country 歌曲風格的歌詞中，「創意與音樂表現」的得分為 0.17、「社交和娛樂活動」的得分為 0.20、「日常和生活情境」的得分為 0.22、「情感和內心世界」的得分為 0.28、以及「感官和身體體驗」的得分為 0.14。其中「日常和生活情境」、以及「創意與音樂表現」的主題分數相對高於整體歌詞(Overall)之平均分數之占比。由此可知，Country 歌曲風格的歌詞內容傾向於將日常和生活融合於音樂表現之中，充滿著對環境周遭與生活情境的描述，例如對大自然的觀察、以及對親友、故鄉的思念之情。透過「創意與音樂表現」主題所對應之候選關鍵字(表 1)，可觀察出該主題中包含 fun(樂趣)、metaphor(隱喻)、rollin(滾動)、singing(唱歌)、 time(時間)、run(奔跑)、ring(鈴響)、raining(下雨)、song(歌)、meal(餐點)、life(生活)、 care(關懷)、以及 chirp(蟲鳥叫聲)等字詞；而「日常和生活情境」主題所對應之候選關鍵字包含 hound(獵犬)、country(鄉下)、steamship(輪船)、house(房子)、man(男人)、work(工作)、dog(狗)、sex(性)、brag(炫耀)、alright(好吧)、virginia(弗吉尼亞州)、fleshy(豐滿)、gather(採集)、tender(溫柔)、sweet(甜)、drunk(醉)、以及 brave(勇氣)等字詞。Country 風格的歌曲最早是起源於 1920 年的美國南方，編曲結構相對於其它歌曲風格，顯得簡潔又明亮，歌詞文本傾向於正面情緒，充滿自由奔放、對自然的探索、以及對未知旅程追尋的景象(Alexopoulos & Taylor 2020)。敘事內容大多集中在創作者的生活環境，呈現了當時純樸與原始的生活型態(Smith 1980; Van Sickel 2005)。例如列舉於圖 10 Country 類別的第 1 首，John Denver 的經典歌曲《Take Me Home, Country Roads》，歌詞中描述旅程、漫遊和回家的道路，述說著在外地遊子期望終有一天能踏上歸鄉旅途的鄉間小路，返回記憶中有山脈、河川、以及採礦生活的故鄉西維吉尼亞(West Virginia)。

Disco 的歌曲有些許風格與 Hip-Hop 相似(圖 5、圖 6)，而其歌詞文字主題傾向亦大部分與 Hip-Hop 相同(圖 9、附圖 A.1)，其中 Disco 的歌詞文字主題相對於 Hip-Hop 更傾向於「情感和內心世界」。Metal 與 Rock 也具備類似的情況，其歌曲風格彼此相似(圖 6)，而兩者的歌詞主題傾向也極為接近(圖 9、附圖 A.1)，這兩類別的歌詞文字主題相較於整體歌詞(Overall)而言更著重於「感官和身體體驗」面向，尤以 Metal 為最。Country 與 Rock，兩者的歌詞主題雷達圖的樣態極為相似(圖 10)，然而 Rock 的「情感和內心世界」與「感官和身體體驗」的傾向表現則更為強烈，相反的 Country 則傾向於「日常和生活情境」，而其歌曲風格之音樂成份亦頗為相似(圖 7、圖 8)。

Blues	創意	社交	日常	情感	感官
第 1 首	0.12	0.23	0.23	0.32	0.10
第 2 首	0.09	0.09	0.38	0.24	0.20
第 3 首	0.14	0.22	0.18	0.37	0.08
第 4 首	0.06	0.15	0.25	0.41	0.13
第 5 首	0.29	0.08	0.22	0.24	0.17
第 6 首	0.13	0.29	0.20	0.27	0.11
第 7 首	0.20	0.28	0.09	0.30	0.12
第 8 首	0.14	0.37	0.26	0.16	0.08
第 9 首	0.24	0.04	0.03	0.67	0.01
第 10 首	0.06	0.08	0.26	0.42	0.18
Mean	0.15	0.18	0.21	0.34	0.12
Min	0.06	0.04	0.03	0.16	0.01
Max.	0.29	0.37	0.38	0.67	0.20
Std.	0.07	0.11	0.09	0.13	0.05

Country	創意	社交	日常	情感	感官
第 1 首	0.06	0.10	0.63	0.10	0.10
第 2 首	0.16	0.18	0.17	0.38	0.11
第 3 首	0.10	0.11	0.14	0.51	0.15
第 4 首	0.09	0.20	0.25	0.15	0.31
第 5 首	0.27	0.40	0.07	0.19	0.06
第 6 首	0.18	0.29	0.07	0.33	0.13
第 7 首	0.23	0.12	0.14	0.33	0.17
第 8 首	0.24	0.13	0.24	0.34	0.05
第 9 首	0.18	0.20	0.21	0.28	0.13
第 10 首	0.20	0.24	0.27	0.13	0.15
Mean	0.17	0.20	0.22	0.28	0.14
Min	0.06	0.10	0.07	0.10	0.05
Max.	0.27	0.40	0.63	0.51	0.31
Std.	0.06	0.09	0.15	0.12	0.07

Rock	創意	社交	日常	情感	感官
第 1 首	0.11	0.14	0.14	0.54	0.07
第 2 首	0.06	0.34	0.13	0.30	0.17
第 3 首	0.04	0.06	0.36	0.39	0.15
第 4 首	0.06	0.27	0.29	0.15	0.23
第 5 首	0.11	0.22	0.16	0.11	0.40
第 6 首	0.13	0.32	0.06	0.09	0.40
第 7 首	0.19	0.07	0.04	0.63	0.08
第 8 首	0.05	0.11	0.14	0.63	0.07
第 9 首	0.20	0.12	0.13	0.44	0.11
第 10 首	0.17	0.13	0.26	0.36	0.09
Mean	0.11	0.18	0.17	0.36	0.18
Min	0.04	0.06	0.04	0.09	0.07
Max.	0.20	0.34	0.36	0.63	0.40
Std.	0.06	0.10	0.10	0.19	0.12

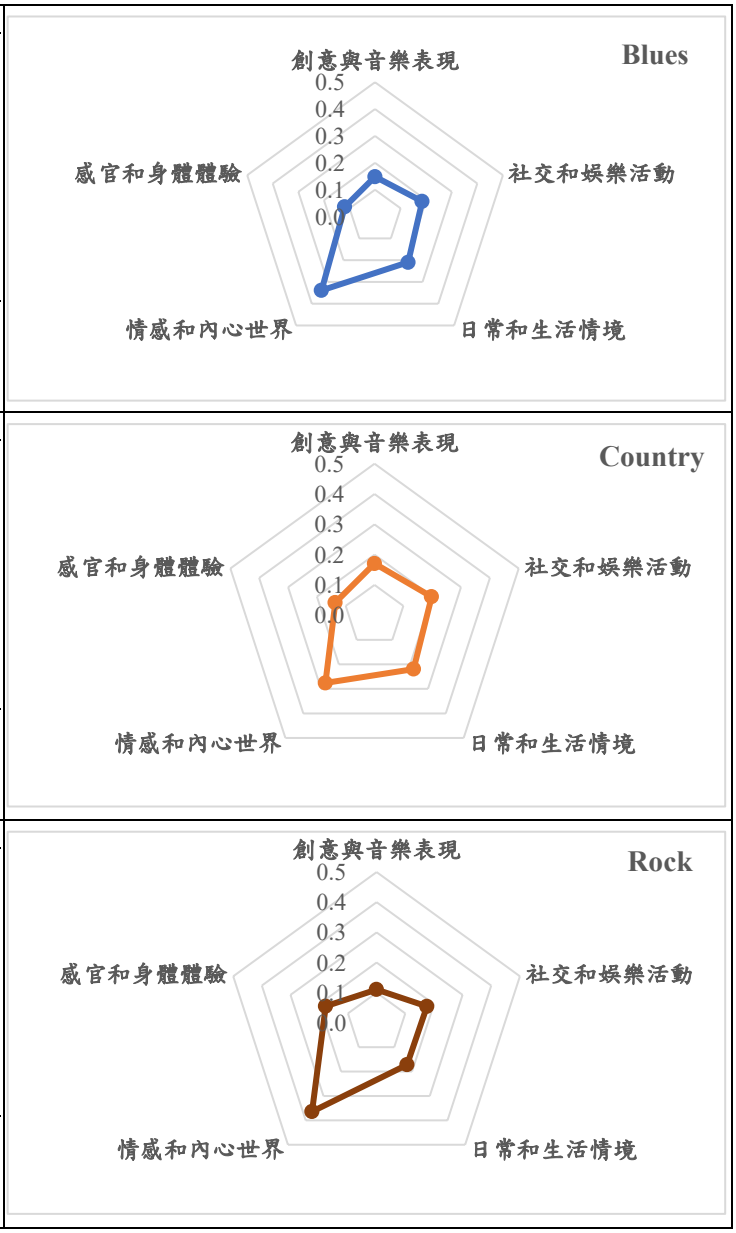


圖 10：Blues, Country, Rock 歌曲類別之歌詞主題傾向雷達圖分析

透過對歌曲風格音樂成份分析，發現歌曲風格的構成元素往往並不單一存在，不同曲風之間常常相互影響。同時，透過歌詞主題傾向雷達圖分析，也發現歌詞中的文字常常抒發著創作者的經歷和心路歷程。將歌曲風格音樂成份分析與歌詞主題傾向雷達圖分析整合起來，可以更全面地理解歌曲與歌詞的內容，證明了歌曲風格與歌詞文字之間的緊密關聯，這種關聯鏈結著創作者的生活經驗、文化背景和歷史脈絡。因此，音樂數位典藏若採用本研究所提出的歌曲風格分類、以及歌詞主題傾向等分析機制來彙整樂曲、樂譜、歌曲、歌詞、以及創作背景等典藏內容，相信對於音樂素材的檢索、以及音樂知識庫的建構將更具效益。

伍、結論

本研究採用卷積神經網路模型（CNN）架構，透過調整多種模型訓練參數、音訊資料增強模式以及梅爾頻譜轉換等策略，以增強歌曲風格分類模型的品質，

並將其應用於歌曲風格音樂成分分析。同時，透過組合式主題模型(CombinedTM)，針對歌詞文字語境進行歌詞主題傾向雷達圖的構建，用於分析歌詞主題傾向。最後，基於歌曲風格分類結果和歌詞主題傾向分析，整合探討歌曲和歌詞之間的關聯性。為了評估方法的可行性，本研究參考 Apple Music 串流平台規範的歌曲風格類別，並隨機選擇該類別歌曲進行歌曲風格分類和歌詞主題傾向的驗證。結果顯示，所構建的歌曲風格分類模型能夠準確分類和分析指定歌曲風格的音樂構成成分組合，分類準確率高達 0.89。在歌詞文字內容的主題傾向分析部分，可透過歌詞主題傾向雷達圖，藉由創意與音樂表現、社交和娛樂活動、日常和生活情境、情感和內心世界、以及感官和身體體驗等方面，解釋各種類別歌詞中的文本語義脈絡。整合歌曲風格分類模型和歌詞主題傾向雷達圖，採用本研究提出的「歌曲類別與歌詞主題傾向雷達圖」進行分析，將清晰揭示音樂和文本之間的相互影響，更有助於我們理解音樂作品的詞曲關係。

然而，實驗所採用的 GTZAN 音樂素材涵蓋了眾多早期作品，代表性可能受限，而且所採用的 Apple Music 平台之作品跨足多個年代，其音樂風格和歌詞亦可能存在時代差異，因而可能影響模型的預測能力，為其限制。在這些方面將需要運用其他方法加以驗證，或採用更多案例研究來證明這些模型在音樂資訊檢索中的實際效果。除此之外，建議未來研究可結合其它技術或領域的應用，例如透過情感分析解析歌詞背後的情感與意義，或是將歌曲風格音樂成分分析和歌詞主題傾向雷達圖分析整合到音樂數位典藏系統或網路音樂串流平台等應用中，以強化音樂知識的建構和音樂內容檢索的效益。

致謝

本研究為國科會計畫(計畫編號：NSTC 112-2410-H-036-002)、教育部教學實踐計畫(計畫編號：PBM1120352)、以及大同大學 USR 深耕型計畫(案號：A8430-1112002)支持補助執行之部分研究成果。

參考文獻

李怡萱 (2019)，「*Ma Rainey 藍調歌曲研究：性別、社會與自由*」，未出版碩士論文，國立臺灣大學音樂學研究所，台北市。

宋英維 (2011)，「古典跨界音樂之重思」，*音樂研究*，頁 25-49。

郝沛毅、龔千芬、張俊陽、蔣榮先、鄭詠恆 (2020)，「嶄新的即時 POI 推薦系統—使用即時事件，圖文/時間內容感知資訊與樹狀卷積網路」，*資訊管理學報*，27(4)，頁 495-535。

黃芝璇、馬麗菁 (2021),「以圖形方式摘要化顧客評論」,*資訊管理學報*,28(2),頁 125-153。

鄭英傑 (2022),「書評:「戰場」轉移—評介《流行樂、媒體與青少年文化:從「節拍革命」到「位元世代」》」,*教育研究集刊*,頁 113-124。

賴靈恩 (2015),「回歸部落之歌:卑南族下賓朗部落的歌謠分類」,*臺灣音樂研究*,頁 97-135。

Aguiar, R. L., Costa, Y. M. G., & Silla, C. N. (2018). *Exploring Data Augmentation to Improve Music Genre Classification with ConvNets.* 2018 *International Joint Conference on Neural Networks (IJCNN)*, Rio de Janeiro, Brazil.

Alexopoulos, C. & Taylor, L. D. (2020). Easy listening? An analysis of infidelity in top pop, hip-hop, and country song lyrics over 25 years. *Psychology of Music*, 48(6), 795-807.

Bahuleyan, H. (2018). Music genre classification using machine learning techniques. *arXiv:1804.01149*.

Bianchi, F., Terragni, S., & Hovy, D. (2020). Pre-training is a hot topic: Contextualized document embeddings improve topic coherence. *arXiv pre-print server*. doi:10.18653/v1/2021.acl-short.96

Bianchi, F., Terragni, S., Hovy, D., Nozza, D., & Fersini, E. (2020). Cross-lingual contextualized topic models with zero-shot learning. *arXiv preprint arXiv:2004.07737*.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research,* 3, 993-1022.

Bromley, J., Guyon, I., LeCun, Y., Säckinger, E., & Shah, R. (1993). Signature verification using a "siamese" time delay neural network. *Advances in Neural Information Processing Systems, 6*.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., & Askell, A. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems,* 33, 1877-1901.

Byrd, D. (2023). "Why I Sing the Blues": The blues and the individuals who played them. *All Theses*, 4070.

Byrd, D. & Crawford, T. (2002). Problems of music information retrieval in the real world. *Information Processing & Management*. 38(2), 249-272.

Churchill, R. & Singh, L. (2022). The evolution of topic modeling. *ACM Computing Surveys*, 54(10s), 1-35.

Compagna, A. (2001). The devil in Robert Johnson: The progression of the delta blues to rock and roll, http://www.loyno.edu/--history/journal/ 1999-2000/Compagna. htm.

Costa, Y., Oliveira, L., Koerich, A., & Gouyon, F. (2013). Music genre recognition based on visual features with dynamic ensemble of classifiers selection. 2013 *20th International Conference on Systems, Signals and Image Processing (IWSSIP),* 55-58.

De Valk, R., Volk, A., Holzapfel, A., Pikrakis, A., Kroher, N., & Six, J. (2017). MIRchiving: Challenges and opportunities of connecting MIR research and digital music archives. *Proceedings of the 4th International Workshop on Digital Libraries for Musicology*, 25-28.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*, Minneapolis, Minnesota.

Dong, Q., Li, L., Dai, D., Zheng, C., Wu, Z., Chang, B., Sui, Z. (2023). A survey on in-context learning. *arXiv pre-print server*. doi:Nonearxiv:2301.00234

Doon, R., Rawat, T. K., & Gautam, S. (2018). Cifar-10 Classification using deep convolutional neural network. 2018 *IEEE Punecon*.

Downie, J. S. (2003). Music information retrieval. *Annual Review of Information Science and Technology*, 37(1), 295-340.

Elbir, A. M. (2020). DeepMUSIC: Multiple signal classification via deep learning. *IEEE Sensors Letters*, 4(4), 1-4.

Fang, J., Grunberg, D., Litman, D. T., & Wang, Y. (2017). Discourse analysis of lyric and lyric-based classification of music. The international society for music information retrieval (ISMIR), 464-471.

Fell, M., Cabrio, E., Tikat, M., Michel, F., Buffa, M., & Gandon, F. (2023). The WASABI song corpus and knowledge graph for music lyrics analysis. *Language Resources and Evaluation*, 57(1), 89-119.

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2022). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *BioRxiv*, 2022-12.

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digital Medicine*, 6(1), 75.

Glorot, X. & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *Proceedings of The Thirteenth International Conference on Artificial Intelligence and Statistics*.

Greene, D., O'Callaghan, D., & Cunningham, P. (2014). How many topics? Stability analysis for topic models. *Machine Learning and Knowledge Discovery in*

Databases: European Conference, ECML PKDD 2014, Nancy, France, September 15-19, 2014. Proceedings, Part I 14 (498-513). Springer Berlin Heidelberg.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proceedings of the IEEE international Conference on Computer Vision*.

Hoffer, E. & Ailon, N. (2015). *Deep Metric Learning Using Triplet Network,* Similarity-based pattern recognition: third international workshop, SIMBAD 2015, Copenhagen, Denmark, October 12-14, 2015.

Hu, X., Downie, J. S., & Ehmann, A. F. (2009). Lyric text mining in music mood classification. *American Music*, 183(5,049), 2-209.

Hwang, M. I. & Lin, J. W. (1999). Information dimension, information overload and decision quality. *Journal of Information Science*, 25(3), 213-218.

Jamdar, A., Abraham, J., Khanna, K., & Dubey, R. (2015). Emotion analysis of songs based on lyrical and audio features. *arXiv preprint arXiv:1506.05012.*

Kingma, D. P. & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278-2324. doi:10.1109/5.726791.

Li, Y., Zhang, Z., Ding, H., & Chang, L. (2023). Music genre classification based on fusing audio and lyric information. *Multimedia Tools and Applications*, 82(13), 20157-20176.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis,* 42, 60-88. doi:10.1016/j.media.2017.07.005.

Loper, E. & Bird, S. (2002). Nltk: The natural language toolkit. *arXiv preprint cs/0205028*.

Martinez, A. M. (2021). Suburban cowboy: Country music, punk, and the struggle over space in orange county, 1978-1981. *California History*, 98(1), 83-97.

Nirmal, M. R. & Mohan B S, S. (2020). Music genre classification using spectrograms. 2020 *International Conference on Power, Instrumentation, Control and Computing (PICC)*, Thrissur, India. https://dx.doi.org/10.1109/picc51425.2020.9362364.

McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O. (2015). Librosa: Audio and music signal analysis in Python. *Proceedings of the 14th Python in Science Conference*.

Meyers, A., Johnston, N., Rathod, V., Korattikara, A., Gorban, A., Silberman, N., & Murphy, K. P. (2015). Im2Calories: towards an automated mobile vision food diary.

Proceedings of the IEEE International Conference on Computer Vision, 1233-1241.

Miao, Y., Yu, L., & Blunsom, P. (2016). *Neural Variational Inference for Text Processing. International Conference on Machine Learning*.

Mulligan, M. (2021). Global music subscriber market shares Q1 2021. *MIDiA Research.* Retrieved from https://www.midiaresearch.com/blog/global-music-subscriber-market-shares-q1-2021.

Mulligan, M. (2022). Music subscriber market shares Q2 2021. *MIDiA Research.* Retrieved from https://www.midiaresearch.com/blog/music-subscriber-market-shares-q2-2021.

Narváez, P. (1994). The influences of Hispanic music cultures on African-American blues musicians. *Black Music Research Journal*, 14(2), 203-224.

Orio, N., Snidaro, L., Canazza, S., & Foresti, G. L. (2009). Methodologies and tools for audio digital archives. *International Journal on Digital Libraries*, 10, 201-220.

Poonia, S., Verma, C., & Malik, N. (2022). Music genre classification using machine learning: *A Comparative Study*, 13, 15-21.

Pruitt, C. (2019). “Boys Round Here”: Masculine life-course narratives in contemporary country music. *Social Sciences*, 8(6), 176.

Qiang, J., Qian, Z., Li, Y., Yuan, Y., & Wu, X. (2020). Short text topic modeling techniques, applications, and performance: a survey. *IEEE Transactions on Knowledge and Data Engineering,* 34(3), 1427-1445.

Raheb, K. E., Kougioumtzian, L., Stergiou, M., Petousi, D., Katifori, A., Servi, K., & Ioannidis, Y. (2022). Designing an augmented experience for a music archive: What does the audience need beyond the sense of hearing? *ACM Journal on Computing and Cultural Heritage*, 15(4), 1-24.

Reimers, N. & Gurevych, I. (2019). *Sentence-BERT: Sentence Embeddings Using Siamese BERT-Networks*, Hong Kong, China.

Rothenbuhler, E. W. (2007). For-the-record aesthetics and robert johnson's blues style as a product of recorded culture. *Popular Music*, 26(1), 65-81.

Sbalchiero, S. & Eder, M. (2020). Topic modeling, long texts and the best number of topics. Some problems and solutions. *Quality & Quantity*, 54, 1095-1108.

Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception*, 7(2), 109-149.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556.*

Singh, Y. & Biswas, A. (2022). Robustness of musical features on deep learning models for music genre classification. *Expert Systems with Applications*, 199, 116879.

Shorten, C. & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data,* 6(1), 1-48.

Smith, S. A. (1980). Sounds of the south: the rhetorical saga of country music lyrics. *Southern Journal of Communication*, 45(2), 164-172.

Srivastava, A. & Sutton, C. (2017). Autoencoding variational inference for topic models. *arXiv preprint arXiv:1703.01488*.

Stevens, S. S., Volkmann, J., & Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *The Journal of the Acoustical Society of America,* 8(3), 185-190.

Stewart, J. B. (2005). Message in the music: Political commentary in black popular music from rhythm and blues to early hip hop. *The Journal of African American History*, 90(3), 196-225.

Sturm, B. L. (2012). An analysis of the GTZAN music genre dataset. *Proceedings of the Second International ACM Workshop on Music Information Retrieval with User-centered and Multimodal Strategies*, 7-12

Tzanetakis, G. & Cook, P. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing,* 10(5), 293-302.

Van Sickel, R. W. (2005). A world without citizenship: On (the absence of) politics and ideology in country music lyrics, 1960-2000. *Popular Music and Society*, 28(3), 313-331.

Vayansky, I. & Kumar, S. A. (2020). A review of topic modeling methods. *Information Systems*, 94, 101582.

Wang, H. C., Syu, S. W., & Wongchaisuwat, P. (2021). A method of music autotagging based on audio and lyrics. *Multimedia Tools and Applications*, 80, 15511-15539.

Wegel, R. L. & Lane, C. E. (1924). The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear. *Physical Review,* 23(2), 266-285. doi:10.1103/PhysRev.23.266

Yu, Y., Luo, S., Liu, S., Qiao, H., Liu, Y., & Feng, L. (2020). Deep attention based music genre classification. *Neurocomputing*, 372, 84-91.

Zhao, W., Chen, J. J., Perkins, R., Liu, Z., Ge, W., Ding, Y., & Zou, W. (2015). A heuristic approach to determine an appropriate number of topics in topic modeling. In BMC bioinformatics. *BioMed Central*, 16, 1-10.

附錄

附表 A.1：歌曲風格分類模型 Model 41 及 Model 45 之精確率(Precision)、召回率(Recall)、以及 F1-score

歌曲風格	Model 41			Model 45		
	Precision	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**
Blues	0.94	0.88	0.91	0.90	0.94	0.92
Classical	0.98	0.96	0.97	0.98	0.94	0.96
Country	0.82	0.66	0.73	0.72	0.78	0.75
Disco	0.87	0.82	0.85	0.85	0.80	0.82
Hip-hop	0.93	0.86	0.90	0.73	0.92	0.81
Metal	0.89	0.96	0.92	0.89	0.94	0.91
Pop	0.84	0.92	0.88	0.98	0.80	0.88
Reggae	0.87	0.94	0.90	0.98	0.80	0.88
Rock	0.72	0.84	0.78	0.73	0.74	0.73

附表 A.2：ChatGPT 提示語

假設你是一位語言專家。請用以上 Group 0 ~ Group 4 五群文字(請參閱表 1），分別給予最適合的五群主題名稱，突顯這五群主題名稱之間的差異。並解釋所賦予之主題名稱的原因。 回答格式如下 ：
Group 0: 主題名稱
Group 1: 主題名稱
Group 2: 主題名稱
Group 3: 主題名稱
Group 4: 主題名稱

ChatGPT Feedback:
Group 0: 創意與音樂表現
這組詞彙包含了與創造、音樂和藝術表現相關的詞語，突顯了創意和音樂性的主題。
Group 1: 社交和娛樂活動
這組詞彙聚焦於社交互動、娛樂活動和人際關係，強調了社交和娛樂方面的主題。
Group 2: 日常和生活情境
這組詞彙關聯到地方性的元素、生活情境和文化，強調了地方性和日常生活。
Group 3: 情感和內心世界
這組詞彙涵蓋了情感、內心世界和個人感受，反映了情感和情緒表達的主題。
Group 4: 感官和身體體驗
這組詞彙突顯了強烈感官體驗、身體感知和情感表達，強調了感官和身體在生活中的重要性。

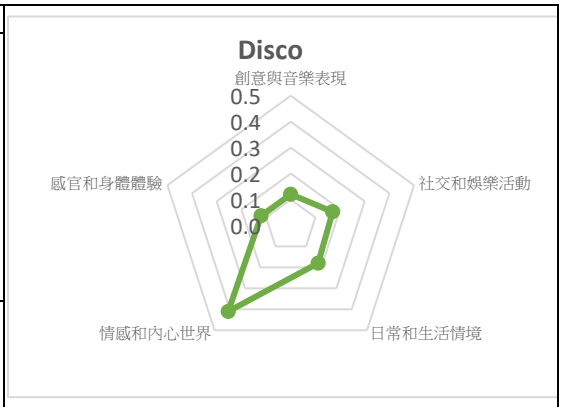
附表 A.3：Apple Music 音樂清單

✓	Black Leaves	✓	The Eternal Sea
✓	Catfish Blues	✓	ye entrancemperium
✓	I Can't Be Satisfied	✓	Oden Owens You All
✓	Smokestack Lightnin	✓	Dismembered
✓	The Cross Road Blues	✓	Angel Of Death
✓	baby what you want me to do	✓	We Eat Our Young
✓	call it stormy monday	✓	Meet Your Maker
✓	Let the Mermaids Flirt with me	✓	Dark Medieval Times
✓	One Bourbon, One Scotch, One Beer	✓	The Old Hag
✓	Ocean of tears	✓	The Wrong Time
✓	Take Me Home, Country Roads	✓	Poker Face
✓	Fifth of May	✓	Shape Of You
✓	Goldmine	✓	Levitating
✓	Can I Take My Hounds to Heaven	✓	Umbrella
✓	Please Don't Go	✓	Call Me Maybe
✓	Inconvenience Store	✓	Always Be My Baby
✓	Stone	✓	Teenage Dream
✓	Wrong Side of the River	✓	Chandelier
✓	No Horse To Ride	✓	Wrecking Ball
✓	Highway Anthem	✓	Moves Like Jagger
✓	Le Freak	✓	Murder She Wrote
✓	Youre Gonna Make Me Love Somebody Else	✓	Big Phat Fish
		✓	Everyone Falls in Love
✓	Make You Cry	✓	No Letting Go
✓	Wild Life	✓	No Games
✓	Funkytown	✓	Heads High
✓	Rasputin	✓	Toast
✓	Ring My Bell	✓	Gimme The Light
✓	Y.M.C.A.	✓	Action
✓	Get Down Tonight	✓	Hold You
✓	September	✓	You Shook Me All Night Long
✓	IG	✓	I Still Haven't Found What I'm Looking For
✓	Water		
✓	Sex On The Moon	✓	Sweet Child O' Mine
✓	We Caa Done	✓	Gimme Shelter
✓	J'adore	✓	Nothing Else Matters (Elevator Version)
✓	Hell Yeah	✓	Smells Like Teen Spirit
✓	Schizo	✓	Another One Bites The Dust
✓	No Time Wasted	✓	Summer Of '69
✓	Tony Soprano 2	✓	Subterranean Homesick Blues
✓	Public Figure	✓	Take It Easy

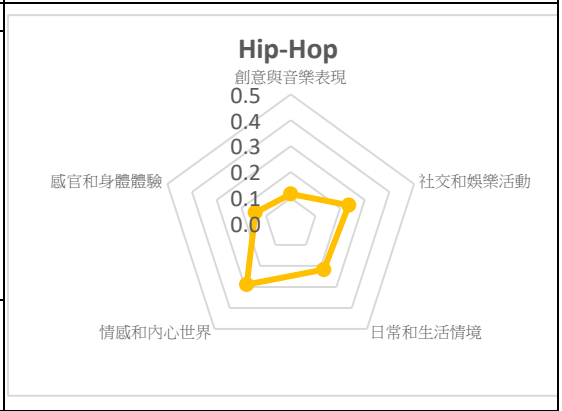
附表 A.4：八種歌曲類別對應整體歌詞(Overall)主題傾向之 Z-分數

歌曲風格	創意與音樂表現	社交和娛樂活動	日常和生活情境	情感和內心世界	感官和身體體驗
Blues	0.811	-0.716	0.836	0.548	-0.606
Country	1.732	-0.486	1.248	-0.439	-0.233
Disco	-0.570	-0.946	-0.355	1.645	-0.650
Hip-Hop	-0.570	0.121	1.156	-0.235	-0.079
Metal	-1.575	0.958	-1.042	-1.645	2.354
Pop	0.644	-0.240	0.515	0.235	-0.584
Reggae	0.351	2.156	-1.454	-1.050	-0.891
Rock	-0.821	-0.847	-0.904	0.940	0.688

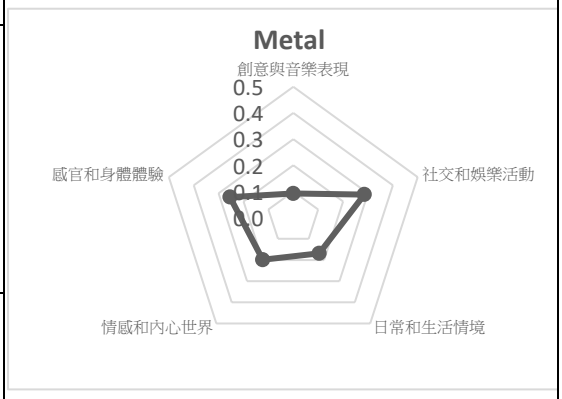
Disco	創意	社交	日常	情感	感官
第 1 首	0.16	0.44	0.06	0.30	0.03
第 2 首	0.17	0.16	0.26	0.24	0.17
第 3 首	0.14	0.28	0.12	0.30	0.16
第 4 首	0.13	0.13	0.33	0.23	0.17
第 5 首	0.02	0.08	0.02	0.87	0.00
第 6 首	0.05	0.04	0.27	0.57	0.07
第 7 首	0.17	0.17	0.03	0.61	0.01
第 8 首	0.09	0.08	0.38	0.15	0.31
第 9 首					
第 10 首	＊Disco 類別只有 8 首歌詞				
Mean	0.12	0.17	0.18	0.41	0.12
Min	0.02	0.04	0.02	0.15	0.00
Max.	0.17	0.44	0.38	0.87	0.31
Std.	0.05	0.12	0.13	0.23	0.10



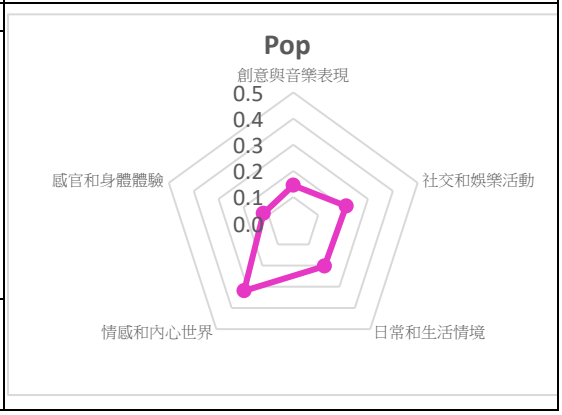
Hip-hop	創意	社交	日常	情感	感官
第 1 首	0.06	0.15	0.23	0.23	0.32
第 2 首	0.15	0.42	0.17	0.21	0.05
第 3 首	0.19	0.15	0.34	0.28	0.05
第 4 首	0.02	0.19	0.24	0.42	0.13
第 5 首	0.08	0.37	0.17	0.26	0.12
第 6 首	0.08	0.43	0.16	0.15	0.18
第 7 首	0.07	0.11	0.13	0.56	0.13
第 8 首	0.09	0.39	0.08	0.30	0.15
第 9 首	0.07	0.13	0.43	0.24	0.14
第 10 首	0.35	0.02	0.22	0.24	0.18
Mean	0.12	0.24	0.22	0.29	0.14
Min	0.02	0.02	0.08	0.15	0.05
Max.	0.35	0.43	0.43	0.56	0.32
Std.	0.09	0.14	0.10	0.11	0.07



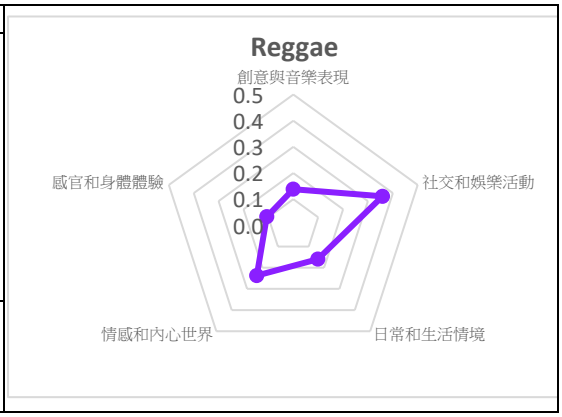
Metal	創意	社交	日常	情感	感官
第 1 首	0.13	0.22	0.09	0.22	0.33
第 2 首	0.06	0.20	0.21	0.07	0.46
第 3 首	0.04	0.23	0.24	0.09	0.40
第 4 首	0.14	0.15	0.08	0.27	0.36
第 5 首	0.03	0.24	0.13	0.40	0.20
第 6 首	0.07	0.50	0.18	0.08	0.16
第 7 首	0.12	0.32	0.11	0.26	0.19
第 8 首	0.09	0.40	0.19	0.23	0.09
第 9 首	0.17	0.18	0.34	0.19	0.13
第 10 首	0.08	0.42	0.11	0.18	0.21
Mean	0.09	0.29	0.17	0.20	0.25
Min	0.03	0.15	0.08	0.07	0.09
Max.	0.17	0.50	0.34	0.40	0.46
Std.	0.04	0.11	0.08	0.10	0.12



Pop	音樂	自然	社會	情感	夢想
第 1 首	0.09	0.42	0.09	0.30	0.10
第 2 首	0.19	0.01	0.77	0.02	0.02
第 3 首	0.18	0.19	0.17	0.31	0.14
第 4 首	0.09	0.26	0.16	0.39	0.11
第 5 首	0.05	0.37	0.04	0.52	0.01
第 6 首	0.24	0.13	0.10	0.20	0.33
第 7 首	0.11	0.28	0.16	0.29	0.17
第 8 首	0.26	0.07	0.07	0.56	0.04
第 9 首	0.10	0.15	0.21	0.31	0.22
第 10 首	0.14	0.24	0.27	0.29	0.07
Mean	0.14	0.21	0.20	0.32	0.12
Min	0.05	0.01	0.04	0.02	0.01
Max.	0.26	0.42	0.77	0.56	0.33
Std.	0.07	0.12	0.20	0.14	0.09



Reggae	音樂	自然	社會	情感	夢想
第 1 首	0.14	0.03	0.45	0.13	0.24
第 2 首	0.33	0.18	0.16	0.25	0.08
第 3 首	0.18	0.60	0.08	0.08	0.07
第 4 首	0.22	0.14	0.16	0.44	0.05
第 5 首	0.11	0.29	0.10	0.30	0.19
第 6 首	0.01	0.90	0.01	0.05	0.04
第 7 首	0.09	0.42	0.25	0.12	0.11
第 8 首	0.17	0.36	0.01	0.22	0.24
第 9 首	0.09	0.10	0.32	0.48	0.01
第 10 首	0.05	0.57	0.06	0.30	0.03
Mean	0.14	0.36	0.16	0.24	0.11
Min	0.01	0.03	0.01	0.05	0.01
Max.	0.33	0.90	0.45	0.48	0.24
Std.	0.09	0.26	0.14	0.14	0.08



附圖 A.1：Disco, Hip-hop, Metal, Pop, Reggae 歌曲類別之歌詞主題傾向雷達圖分析