

# 應用網路聲量及情緒分析於熱門歌曲點播量預測

陳震瑜

中央大學資訊工程  
107522103@cc.ncu.edu.tw

張嘉惠

中央大學資訊工程  
chia@csie.ncu.edu.tw

邱裕明

KKlab Technologies Limited  
ardenchiu@kklab.com

**Abstract**—社群網站的平台提供群眾發聲的管道，使得大量豐富且多元的意見可以被聽見及討論，也讓輿情分析變成新興的市場調查方法。對於音樂串流平台來說，社群網路之輿情分析有助於解決客群侷限於特定平台的問題，使平台能夠獲得較為全面性的數據來進行分析與決策。本研究運用輿情分析結果，鎖定在華語歌曲熱門歌曲點播量預測。我們分析來自於「批踢踢 PTT 實業坊」的社群媒體上與歌手、歌曲的網路聲量，並採用意見目標的情緒分類 Aspect-based sentiment analysis (ABSA) 模型，來對與歌手、歌曲相關的評論進行意見目標之情感分析，將結果應用於「熱門歌曲點播率之預測」。實驗結果顯示，加入「歌手網路聲量」以及「意見目標之情感」分析結果，有助於後續歌曲點播量提升的預測，而且加入「意見目標之情感」分析結果，比起只包含意見目標的「句子整體情感」分析效能，對「熱門歌曲點播率之預測」更為準確。

**Index Terms**—深度學習，輿情分析，熱門歌曲預測

## I. INTRODUCTION

現今網際網路的蓬勃發展，各種影音串流平台、網路論壇等網路服務因應而起，平台為了與使用者交流，設置留言區或是收藏功能等的反饋機制，讓這些社群媒體中產生了多樣化的豐富資料。使網路輿情分析 (Opinion Analysis) 的發展極具潛力，成為近年來熱門的研究議題。在商業市場上，企業能夠透過網路輿情分析掌握時代趨勢，了解消費者所關心、討論的主題，以及對於企業的服務、產品的正負面評價，並藉由這些資訊調整行銷策略、改進產品生產方向與服務品質；進一步可透過輿情風向創造新的話題，吸引消費者與凝聚更多人氣。

以唱片業、娛樂事業為例，過去挖掘新人常仰賴經紀人的經驗，而今新人發片、簽唱會、代言等所帶來的市場反應，可以透過社群網路意見探勘掌握市場動向，不僅可以幫助其經紀公司記錄旗下藝人發展，也能夠更進一步做為產業行銷與決策的考量。

舉例而言，音樂串流平台雖有豐富的統計資訊，如：歌曲的點擊次數、點擊次數的成長趨勢、以及播放媒體上喜歡和不喜歡的人數等等數據，可用來進行熱門歌曲預測 (Hit Song Prediction)。然而，這些數據的來源往往侷限於使用特定播放媒體的客群，沒辦法獲得較全面性的數據。因此，本研究希望利用社群媒體「批踢踢 (PTT)」上的資料來擴大取得更多元化的市場資訊，解決客群被侷限的問題。具體來說，我們的目的在應用輿情分析技術將文字訊息有效的轉化成數值資料，改善熱門歌曲點播率預測之研究。

我們鎖定在華語區的歌手、歌曲，並將問題分成兩項來探討，第一項是根據 KKBOX APP 平台提供的歌曲發行後 14 天內產生的相關歌曲點播數據做為主要觀測數據，運用此數據預測在未來 15 至 30 天中歌曲的點播量。第二項，根據新歌在發行前 30 天至發行後 14 天內在 KKBOX，收集社群平台上與歌曲、歌手相關的貼文作為輔助數據，運用網路聲量分析、及 Aspect Based

Sentiment Analysis 情感分析所得的資訊，將分析結果進行數值化，來檢視網路聲量、情感分析是否有助於提升歌曲點播量的預測準確率。

本研究中的熱門歌曲點播率預測是屬於迴歸問題，在模型的設計上，參考了與 Breuel 等人 [1] 所提出 AutoMLP 模型的概念，來設計本研究的模型架構，依照不同的輸入大小的特徵序列，調整隱藏層的神經元個數，再透過驗證資料集選擇最佳隱藏層數，讓神經網路找出最佳的網路架構，省去人工處理特徵的過程。實驗結果顯示，AutoMLP 比起統計預測模型 Gradient Boosting Machine (GBM) [5]，可以達到更準確的預測，並對於加入的網路提及聲量。

## II. RELATED WORK

熱門歌曲預測 (Hit Song Prediction) 的目的是預測歌曲是否能夠成為未來熱門歌曲。早期的熱門歌曲預測任務，多數都是使用影音串流平台所提供音頻特徵訊息 (如：每分鐘的節拍、響度等)、歌詞等透過機器學的方法來進行分析，如 Fan 等人是運用 Time Weighted Linear Regression、Support Vector Machine (SVM) [4]，而 kai 等人則是運用 Expectation Maximization [6] 等多項方法來對音頻資料進行解析；而在深度學習的方法上，多數的論文都是使用 CNN 來對音頻資料進行特徵擷取 [12], [13]。上述熱門歌曲預測的方法主要專注於分析音頻特徵，本研究的熱門歌曲資料則是平台上歌曲播放的統計資料，並不包含音頻訊息的迴歸問題。

應用意見分析於迴歸預測的研究包括應用 Twitter 情緒分析於美股漲跌趨勢之預測 [9]，又如透過事件歷史的討論程度、事件的影響力等數據，來進行未來目標的預測 [11]。Hossain 等人 [10] 透過時間軸上的一個平移滑動窗口，收集窗口內每分鐘的數據交易特徵資訊 (如：交易日期、股價、股票的售出货量) 做為輸入資料，並運用 CNN、LSTM、RNN 的深度學習模型，學習每分鐘的特徵之間的相互關係，並預測未來某個時間區段的股價。Selvin 等人 [7] 則是混和 LTSM 與 GRU 的方式，透過大量的每日交易資料 (如：開盤價、收盤價、交易量) 來預測隔天的股價。

## III. METHODOLOGY

### A. 網路提及聲量資料準備

本研究使用的網路聲量資料是來自「批踢踢 PTT 實業坊」與音樂相關討論版的貼文，並以歌曲的歌手名作為關鍵字，搜尋歌曲發行日前 30 與後 14 天之之間與歌手相關的貼文 (圖 1)。上述的搜尋關鍵字不採用歌曲名稱來單獨搜尋，原因在於歌曲名稱的用詞常與一些日常用語相似，造成搜尋出許多的雜訊。每則貼文還包含了留言、推噓數等資訊內容，而推噓代表留言者是否贊同貼文本身。

在文本的處理上，會先清理無意義字元，以及非中文用語

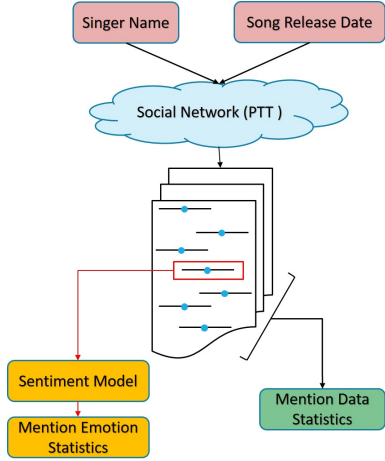


圖 1. 網路聲量資料準備方法

的貼文。接著，運用搜尋出該貼文的歌手名，對貼文進行文本分割，以歌手名為中心抓取上下文各 150 字，經由切割後的文本會產生好幾段包含歌手名的段落，再透過 NTUSD 中情感詞彙字典 [8]，來評估每個段落的情感字詞含量，並找出情感字詞含量較多的段落，應用 BERT-Attention-over-Attention (BERT-AOA) 模型架構進行情感預測 [2]。

在歌曲點播率的預測上，我們比較兩種特徵準備方法，分別為「網路提及聲量 (Network Mention Information)」與「網路提及情感 (Network Mention Emotion)」。

#### 網路提及聲量 (Network Mention Information)

PTT 的貼文除了本文外，也包含了推數、噓數等訊息。本論文規劃出了 4 種方法來統計上述的訊息。

- 1) Mention Per day: 根據歌曲的發行日 (release date) 的前 30 天與後的 14 天統計該歌曲的歌手每日被討論到的文章數量。
- 2) Mention Trend: 以每日討論量為基礎，計算每日討論量的成長趨勢。
- 3) Singer Mention Aggregation: 統計每首歌曲發行日前 30 天與後 14 天，與歌曲的歌手相關貼文的總貼文量、總推數、總噓數。
- 4) Song & Singer Mention Aggregation: 統計統計每首歌曲發行日前 30 天與後 14 天，同時存在含有歌曲和歌手名稱的總貼文量、總推數、總噓數，並結合前一項 Singer Mention Aggregation。

#### 網路提及情感 (Network Mention Emotion)

本研究根據意見目標情感分類模型預測的結果，規劃出 3 項特徵準備方法如下。

- 1) Emotion Per day: 計算每首歌曲發行日前 30 天與後 14 天中，每日在 PTT 上被提及的貼文情感分數。在計算分數上，為了突顯三類情感分數的差異以及易於特徵處理，我們給予三類情感不同的分數，正向->3 分、負向->1 分、中立->2 分。
- 2) 3-Emotion: 統計每首歌曲發行日前 30 天與後 14 天，與歌曲的歌手相關貼文的「正向、負向、中立」三類情感比例。

- 3) 4-Emotion: 以上一個項目為基礎，新增「正向、負向、中立」情感比例中比例最大者，作為加入的新特徵。

#### B. AutoMLP 模型架構

由於本研究後續會加入多種不同維度大小的網路聲量及與情特徵於熱門歌曲預測的特徵集中，形成多種不同的特徵集組合，而不同的特徵集會擁有不同的參數組合，因此我們希望模型可以應付多樣化的特徵輸入。本研究所使用的 AutoMLP 模型設計上是固定學習率，在隱藏層的神經元設計上，根據不同維度大小的輸入特徵，調整隱藏層數與不同的隱藏神經元數量，讓每一層神經元能夠有效的學習特徵之間的關係。而模型的層數也是影響效能的關鍵，因此模型會根據不同評估指標，來選擇每個輸入資料集的適合的隱藏層數。模型架構共有三層，如圖 2。

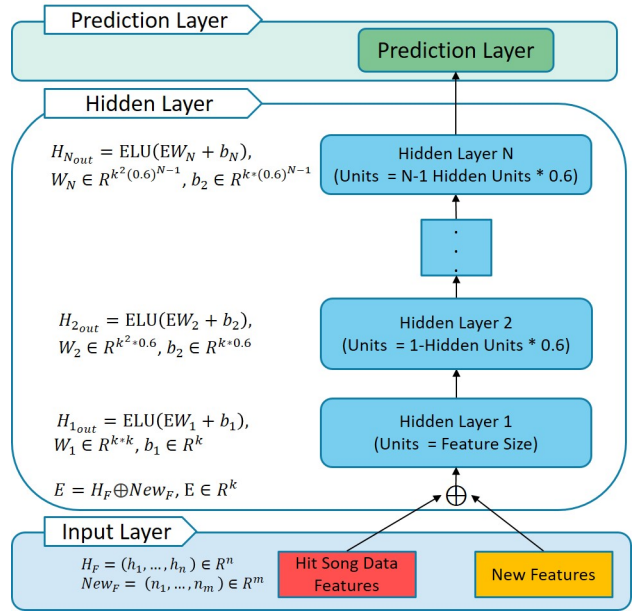


圖 2. AutoMLP 模型架構圖

#### 輸入層 (Input Layer)

熱門歌曲預測特徵是屬於固定特徵維度大小，符號表示為  $H_F = (h_1, \dots, h_n) \in R^n$ ，而新增的網路聲量特徵，根據不同的統計方式，會有不同的特徵維度大小，符號表示為  $New_F = (n_1, \dots, n_m) \in R^m$ ，在此層會將兩者進行進行連接 (connect)，連結後的向量即為輸入神經數量，表示為  $H_0 \in R^{n+m}$ 。

**動態隱藏層 (Dynamic Hidden Layer)** 假設第  $k$  輸出層的特徵維度大小為  $d_k$ ，相當於第  $k+1$  層的輸入維度，則第  $k$  層的輸出如公式 1 所示。

$$H_k = ELU(W_k * H_{k-1} + b_k) \quad (1)$$

$$W_k \in R^{d_k \times d_{k-1}}, b_k \in R^{d_k}$$

其中  $d_0 = m + n$ ，且  $d_k = \lfloor 0.6 * d_{k-1} \rfloor$ 。

**預測層 (Prediction Layer)** 根據最後一個隱藏層的輸出維度  $H_K$  做為輸入，預測熱門歌曲點播量  $\hat{y}$ 。

$$\hat{o} = ELU(W_O * H_K + b_O) \quad (2)$$

$$W_O \in R^{d_K}, b_O \in R$$

此模型的設計與 Breuel 等人 [1] 所提出 AutoMLP 模型的概念相近，都是運用多層的感知器 (Multi-layer Perception, MLP) 架構。Breuel 等人的 AutoMLP 透過驗證集的錯誤率，讓模型在訓練過程中，調整學習率 (learning Rate) 以及隱藏神經元數。我們則著重在層數的調整。

#### IV. Experiment

##### A. 實驗資料集

本研究使用由 KKBOX 提供的影音串流平台資料，在 2019 年 4 月前總共擁有 86,168 首歌曲，而 2019/4 至 2019/7 間則有 4,297 首歌曲。我們挑選在 PTT 中有討論聲量的歌曲，訓練資料中僅有 17,459 首 (20%)，測試資料則是有僅有 1,014 首 (23.5%)，詳細統計資訊如表 I。

表 I  
歌曲點播量預測資料集

Dataset	Training Data	Testing Data
Time Range	Before 2019-04	2019-04 2019-07
# Songs	86,168 (100%)	4,297 (100%)
PTT	17,459 (20.0%)	1,014 (23.5%)

##### B. 評估方法

對於預測歌曲點播量的問題，本論文採用三種常見的指標評估效能：RMSE(Root Mean Square Error)、MAE(Mean Absolute Error) 與 MAPE(Mean Absolute Percentage Error) 評估模型的效能。實務上，公司或投資者在進行投資時，會選擇較熱門與具有潛力的歌曲或歌手。然而，通常熱門的歌曲、歌手，在串流影音平台的點播量會遠高於冷門的歌手或歌曲。當模型對於高點播量的歌曲、歌手預測較差時，相較另外兩個指標，RMSE 會放大錯誤值。反之，預測較好時就相對會低許多。基於熱門歌曲效益對企業更為重要，因此本研究後續實驗結果分析將會著重於 RMSE 指標。

##### C. 實驗結果

由於論文版面的問題，本論文將焦點放在被 PTT 社群媒體上所提及的歌曲進行實驗與分析，並為這些歌曲準備網路聲量特徵，來觀測加入網路聲量特徵是否有助於這些被提及的歌曲在點播量上的預測。透過第三小節中所提到的「網路提及聲量 (Network Mention Information)」與「網路提及情感 (Network Mention Emotion)」，來將兩種不同的特徵準備方法加入熱門歌曲預測的特徵中。

##### 網路提及聲量之預測效能

如表 II 所示。我們以未加入網路提及聲量的原始特徵數據集「Original Features」做為 Baseline (第一列)。整體說來加入不同的網路提及聲量特徵後，有助於改善 GBM 與 AutoMLP 的模型效能，且 AutoMLP 在整體模型效能的表現上都優於 GBM 模型。加入少數網路提及聲量統計值特徵，比起加入 45 天的網路聲量或趨勢，效果更為顯著。對於 GBM 而言，加入過多特徵反而減弱模型效能。

AutoMLP 模型是透過 RMSE、MAE 以及 MAPE 等三項評估指標，來選擇隱藏層數，首先模型會選擇 RMSE 最小的前三名的隱藏層，並透過 MAE 與 MAPE 進一步選擇適合的隱藏層，雖然本研究的分析著重 RMSE 指標，但 MAE 與 MAPE 能夠輔助判斷預測模型的品質。根據

表 II  
網路提及聲量特徵之預測效能

Input	# Features	Evaluation Function	
		RMSE	
		GBM	AutoMLP
Original Features	374	16,559	16,083
Mentions Per day	+45	16,531 (0%)	16,064 (0%)
Mention Trend	+45	17,668 (7%)	15,596 (-3%)
Singer Aggregation	+3	16,522 (0%)	16,275 (1%)
Song & Singer Aggregation	+6	16,203 (-2%)	15,461 (-4%)

圖 3 所示，左子圖代表加入數量較少的聲量特徵，右子圖則代表數量較多的網路提及聲量特徵，根據圖中能觀察到，加入不同維度大小的聲量特徵於原始特徵中，模型都能在 8 到 10 層時獲得最好的效能。

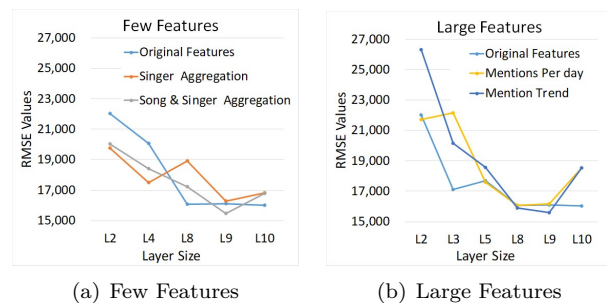


圖 3. 網路提及情感與隱藏層數之效能關係

我們透過預測值與真實值之間的資料分佈圖，來進一步探討加入網路提及聲量的效果，如圖 4 所示，圖中 X 軸為預測值，而 Y 軸為真實值，而紅色代表表現最好的聲量特徵「Song & Singer Aggregation」的預測值，藍色則是未加入任何聲量特徵的「Original Features」的預測值。根據圖 4 的顯示結果，不管在高播放量與低播放量的分布圖中，紅色的資料分佈，相較藍色的資料分佈更集中於中心線，代表加入聲量特徵後，助於提升模型的預測效能。

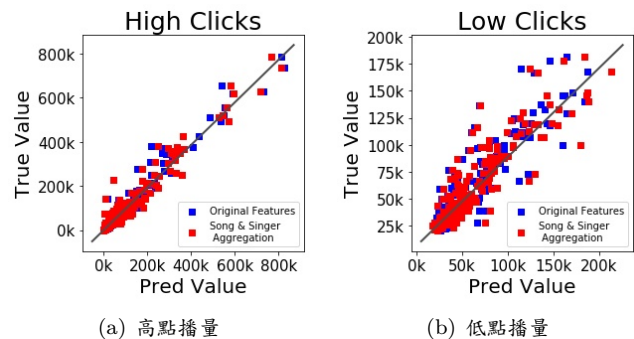


圖 4. 網路提及聲量資料分佈

##### 網路提及情感之預測效能

其次，我們加入兩種不同的情感分析結果於歌曲點播率預測，並比較兩種方法對於預測效果的影響。1. 僅針對包含歌手/歌曲的句子做整體評論情感評價的情感分析。分析方法是參考 [3] 的 BERT 情感模型。2. 針對包含歌手/

表 III  
混和聲量之預測效能

Input	# Features	Evaluation Function					
		RMSE		MAPE		MAE	
		GBM	AutoMLP	GBM	AutoMLP	GBM	AutoMLP
Original Features	374	16,559	16,083	28.1 %	29.0 %	5,936	5,635
4-Emotion	+4	16,073 (-3%)	15,587 (-3%)	28.2 %	28.9 %	5,859 (-1.2%)	5,701 (1%)
Song & Singer Aggregation	+6	16,203 (-2%)	15,461 (-4%)	28.2 %	29.8 %	5,867 (-1.1%)	5,757 (2%)
4-Emotion + singer Aggregation	+7	16,073 (-3%)	15,135(-6%)	28.0 %	28.5 %	5,859 (-1.2%)	5,641 (0.1%)

歌曲的句子進行意見目標情感分析。分析方法是參考 [2] 中所提出的 BERT-AOA 意見目標情感分類模型。

實驗結果如表 IV 所示，兩種情感分析方法，以「4-Emotion」的資料輸入方式更能夠有效改善 GBM 與 AutoMLP 模型預測的效能，且 AutoMLP 在效能表現上優於 GBM 模型，改善幅度也大於 GBM 模型。而運用意見目標情感分析的結果，相較於非目標的情感輿情分析，能夠在預測準確率獲得較大幅度的改善。

表 IV  
網路提及情感聲量之預測效能

Input Features		# Features	Evaluation Function	
			RMSE	
			GBM	AutoMLP
Original Features		374	16,559	16,083
Sentiment Analysis	Emotion Per day	+45	16,410(-0.8%)	15,997 (-0.5%)
	3-emotion	+3	16,480(-0.4%)	16,262 (+1%)
	4-Emotion	+4	16,404(-0.9%)	15,708 (-2.3%)
Aspect Sentiment Analysis	Emotion Per day	+45	16,908 (2%)	15,958 (-0.7%)
	3-emotion	+3	16,321 (-1%)	16,382 (+2%)
	4-Emotion	+4	16,073 (-3%)	15,587 (-3%)

最後我們也嘗試結合「網路提及聲量」與「網路提及情感」將上述兩種不同的聲量特徵同時加入輸入特徵。根據表 III 的實驗結果，混和聲量「4-Emotion + singer Aggregation」預測的準確度上比個別加入「網路提及聲量」或「網路提及情感」的效能更優異 (RMSE 減少 6%)。最後，我們也評估 MAPE 與 MAE 效能，發現混和聲量「4-Emotion + singer Aggregation」僅對 AutoMLP 模型的 MAPE，以及 GBM 模型的 MAE 較有改進空間。其餘兩個聲量「4-Emotion」及「Song & Singer Aggregation」也僅對 GBM 模型的 MAE 效能有所改進。

## V. Conclusion

本研究應用資訊檢索方式取得歌手網路聲量，並將提及歌手、歌曲的相關文章進行意見目標情感分析，得到歌曲之正負中立評價比例。我們比較數種網路提及聲量及情感分析結果的輸入格式，透過自動多層感知機 AutoMLP 模型來進行歌曲點播量預測。實驗結果顯示，AutoMLP 模型的預測效能比統計模型 GBM 預測效果來的更加出色，且增加網路提及聲量特徵，能夠有效的提升熱門歌曲預測的準確率。

本研究結果可作為唱片業在進行發掘熱門歌曲、歌手時的參考資訊。透過社群網路的意見探勘可有效取得歌手、

歌曲之網路聲量，以及對於歌手、歌曲相關評論文章的情感分析，提升歌曲點播率預測的準確率。往後的研究方向，能夠朝向特徵解釋性的相關研究，並擷取出影響熱門歌曲預測的特徵資訊，讓合作商更進一步的分析。

## REFERENCES

- [1] Thomas M. Breuel and Faisal Shafait. Automlpl: simple, effective, fully automated learning rate and size adjustment. In *The learning workshop*, Utah, 2010.
- [2] Chen-Yu Chen and Chia-Hui Chang. 基於 bert-daoa 的意見目標情感分析. In *Proceedings of the 32nd annual Conference on Computational Linguistics and Speech Processing*, September 2020.
- [3] Chen-Yu Chen, Wei-Cheng Chiu, Chia-Hui Chang, Yuh-Ming Chiu, and Hsiu-Min Chuang. 運用多領域資料結合深度學習技術於音樂論壇評論之情感分析. In *2019 International Conference on Technologies and Applications of Artificial Intelligence*, November 2019.
- [4] Jianyu Fan and Michael A. Casey. Study of chinese and uk hit songs prediction. 2013.
- [5] Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29:1189–1232, 2000.
- [6] Dorien Herremans, David Martens, and Kenneth Sørensen. Dance hit song prediction. *Journal of New Music Research*, 43(3):291–302, Jul 2014.
- [7] M. A. Hossain, R. Karim, R. Thulasiram, N. D. B. Bruce, and Y. Wang. Hybrid deep learning model for stock price prediction. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1837–1844, 2018.
- [8] Lun-Wei Ku and Hsin-Hsi Chen. Mining opinions from the web: Beyond relevance retrieval. *Journal of the American Society for Information Science and Technology*, 58(12):1838–1850, 2007.
- [9] V. S. Pagolu, K. N. Reddy, G. Panda, and B. Majhi. Sentiment analysis of twitter data for predicting stock market movements. In *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*, pages 1345–1350, 2016.
- [10] S. Selvin, R. Vinayakumar, E. A. Gopalakrishnan, V. K. Menon, and K. P. Soman. Stock price prediction using lstm, rnn and cnn-sliding window model. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 1643–1647, 2017.
- [11] Yumo Xu and Shay B. Cohen. Stock movement prediction from tweets and historical prices. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1970–1979, Melbourne, Australia, July 2018. Association for Computational Linguistics.
- [12] Li-Chia Yang, Szu-Yu Chou, Jen-Yu Liu, Yi-Hsuan Yang, and Yi-An Chen. Revisiting the problem of audio-based hit song prediction using convolutional neural networks. *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar 2017.
- [13] Lang-Chi Yu, Yi-Hsuan Yang, Yun-Ning Hung, and Yi-An Chen. Hit song prediction for pop music by siamese cnn with ranking loss, 2017.