GET1030

Computers and the Humanities

Tutorial D1

Name	**Matric No.**
Haziq Hakim Bin Abdul Rahman	A0216481H
Ng Jong Ray, Edward	A0216695U
Ong En Jing, Crystal	A0222328M
Rakin Bin Rostam	A0255083E
Zhu Jia Hui, Grace	A0240003H

Introduction

This project explores a comparative analysis of popular TikTok music against popular charting music and the relationship between virality and emotions of music. TikTok is a social media platform where short-form video content is posted and shared among users, and has significantly risen in popularity since 2018 (Eichler, n.d.). The music used in the background of the videos tends to be a snippet of a song – usually the most catchy part of the song like the chorus. TikTok has become an influential marketing tool for artists as the music used in trending videos tend to go viral as well, allowing the musical artists to influence the popularity of their music and its chart performance. This has led to the music industry witnessing a transformative phenomenon that can be coined the "TikTokification" of music: the “tendency to write songs with the specific aim of going viral on TikTok” (Niblett, 2023). It has been shown that the emotions of the music has a large impact on its virality, hence our focus for this project would be to explore the correlation between emotions and virality, identifying the emotions that viral music are generally more inclined towards.

Literature Review

1. Music and TikTok’s impact on recent music

According to Saquete et al. (2022), various factors contribute to virality, including the emotional aspects of content and its intensity. The shareability of content is predominantly influenced by its emotional aspects, where emotions such as happiness, sadness, anger possess the potential to resonate with the audience, thereby enhancing the content's shareability within individual networks and ultimately leading to virality. A study conducted by Saquete et al. (2022) illustrates media with a notably high emotional charge are more likely to be widely circulated and achieve viral status. Due to the high engagement and viewership by users, it was easy for content creators to go viral for their videos or sounds, giving music artists opportunities to gain popularity by making their music viral on TikTok.

TikTok also serves as a marketing tool for artists to increase visibility to their music and accessibility to their fans as well. The platform is responsible for music discovery and breaking new musical acts, and artists have started to take TikTok into consideration when discussing their album marketing strategy. Such tactics have been used in order to achieve virality of music, and it is shown that viral videos using snippets of songs have a direct correlation to big music charting companies like Billboard chart positions (Coulter, 2022).

2. *Informing our research methodology*

Music has been very closely linked to emotions, as one of the main motivations for listening to music is the pursuit of an emotional experience (Juslin & Laukka, 2004). The relationship between music and emotions has been increasingly studied by academic scholars in recent decades.

On research methodology that scholars have used to analyse music, emotion analysis has been a prominent way used to analyse songs and their lyrical and audio features. Jamdar et al. (2015) elucidates both lyrical and audio features are used in tandem as these are crucial musical aspects that determine the emotion of a song: people connect with their music with its lyrics, as words clearly express the emotion of a song; and audio features tie in together with the lyrics as the audio is made to fit with the lyrical theme. Focusing on the lyrical features vis-a-vis the scope of our project's topic, both Jamdar et al. (2015) and Xu et al. (2021) concur that Natural Learning Processing (NLP) technology has been widely adopted by researchers to study the perceived emotions in texts. However, Jamdar et al. (2015) clarify that mere text classification (usually used in the analyses of text documents like news articles, movie reviews) cannot be directly utilised to classify song lyrics, as song lyrics require linguistic principles such as the semantic similarities between words on lyrics in order to capture the meaning behind the lyrics – before attaching a certain emotion to the lyric features.

Similarly, Xu et al. (2021) have utilised "linguistic inquiry and word count technology (LIWC)" to extract lyrical features of Chinese songs, thereafter investigating its effects on the perceived arousal and valence of music. Some of the variables Xu et al. have picked out using LIWC are the total number of lyric words, and proportion of negative emotion words – similar to text classification of data. In addition, they study audio and lyric features using machine learning (ML) models. While we do not build a ML model from scratch, we use '*pysentimentio*' which has already been fed data on music features and its perceived emotions.

Another related work is Napier & Shamir (2018)'s, which compares the change of lyrical tone over a period of 66 years to understand the trend of tones in popular music. Their work used Tone Analyser software, a blend of psycholinguistics and ML to identify and characterise the tone present in a given chorus of lyrics, thereafter attaching an emotion. With this tool, they have analysed the lyrical tone of songs and identified a tonal change over 66 years. Their research provides new insights in the relationship between lyrical tones and emotions, and its historical evolution in popular music.

Methodology

Our methodology section breaks down the steps in sequential order that our group has taken to analyse the data in our comparative analysis of the emotions of lyrics from the most popular songs between TikTok and Billboard:

1. Choosing the lists of popular music

We have extracted data on the top popular songs in the US via the *'TikTok Billboard Top 50*' list (Billboard, n.d.) and the *'Billboard Hot 100'* list (Billboard, n.d.). To create a fair comparison between the two datasets and equalise the number of songs compared on both sides, only the first 50 songs of the *Billboard Hot 100* list have been taken into consideration. Moreover, since the charts are updated every week, we ensured to extract both lists from the same week of 28 October 2023.

We have decided to extract the first 50 songs from the *Billboard Hot 100* list as a proxy for the top 50 most popular music in the US context. The *Billboard Hot 100* is a weekly chart published by the entertainment magazine Billboard since 1958. It is widely recognised and used in pop music studies to analyse popular songs, trends, and preferences. It serves as a key tool for identifying the most popular music hits each year (Napier & Shamir, 2018), as Billboard charts are compiled using information from radio stations in more than 140 markets across the U.S. (Billboard, n.d.). This has informed our decision to utilise the *TikTok Billboard Top 50* list as a proxy for the US's most popular TikTok music. Doing so will standardise the source of the database – that is Billboard – in order to have a comparable set of data to analyse, since Billboard's method of classifying the top 50 songs will be consistent for both lists.

2. Finding lyrics

Contrary to the aforementioned work of Napier & Shamir (2018), we have decided to use a single source to obtain the lyrics of the songs instead of from several sources, due to the scope of our project, as well as to avoid potential discrepancies between lyrics from different sources. We have selected our lyric source to be Genius.com. Based in the US which complements our usage of Billboard charts as aforementioned, Genius is a huge database for song lyrics that are backed up by a community of more than two million contributors from across the globe (Genius, 2016). Lyric datasets are verified by moderators and sometimes by the artists themselves (Genius, n.d.). This database includes the translation of lyrics for non-English songs, which is crucial for our analysis as the Billboard charts we have selected encompasses non-English songs. The platform's extensive database, user verification, and artist validations provides a level of accuracy where Genius's provision of context-sensitive English translations adds another layer of reliability compared to direct translations using tools like Google Translate (Genius, 2017).

2.1. Processing lyrics

For each song-artist pair from both Billboard top 50 lists, we will search using Genius API (Application Programming Interface) to find Genius.com URL (Uniform Resource Locator) link of the song by searching via the song title and artist name. Since the returned URL link will contain both the song title and artist name, this was used to verify if the search has found the correct song – a needed step since many songs exist with the same name.

Since Genius API only returns metadata of the song, web scraping using the Genius URL link was done to find the lyrics. This was done using both Python '*Requests*' and '*BeautifulSoup'* libraries, with inspiration from Khan (Khan, 2021). '*Requests*' was used to get the HTML

(HyperText Markup Language) file and '*BeautifulSoup*' would be used to parse the HTML file and retrieve the lyrics as text. Thereafter, we separated the song lyrics by their different sections such as verse, chorus, bridge and others.

3. Conducting Emotional Analysis

We will be conducting an emotion analysis, which is under text classification and stems from the use of the python library '*Pysentimiento*'. It should also be noted that as per our research's scope, we have eliminated songs without lyrics. More specifically, we will be analysing the emotion that the first chorus of each song brings about.

Firstly, we have decided to use only the first chorus of each song, as choruses are usually described as the section of a song that is more outstanding, catchy, memorable (Balen et al., n.d.), which can help contribute to the virality of TikTok music as per our group's topic. Thereafter, we have eliminated songs that do not have choruses, such as some rap music. We have acknowledged that the selection of chorus limits our emotion analysis because songs are versatile and can present themselves in various musical structures. Additionally, this means excluding rap music despite it being a dominant genre in the US. Like many text analysing tools, focusing solely on the chorus often overlooks the broader context of the lyrical meaning the artists are portraying and the intended emotions (Napier & Shamir, 2018). However, this is a strategic move to ensure consistency across all song lyrics selected, as we were previously unable to come into a consensus as to which section of a rap song should be chosen. Thus, this leads to a total of 42 songs for TikTok and 47 songs for Billboard.

Secondly, we have decided to use six emotions for our emotional analysis: Happiness, Sadness, Anger, Fear, Disgust and Surprise, based on information from Hugging Face (Lim, 2023), an online collaboration platform for ML tools. These emotions have been adapted from American psychologist Ekman (1992) who proposes six primary and universal human emotions. Only six main emotions are chosen due to the scope of our project since we are only using a dataset of top 50 music. An extensive range of emotion categories will prevent meaningful analysis of our relatively small dataset of two lists with 50 songs. For example, we have selected 'Happiness' to be a category on its own, instead of splitting happiness into more nuanced forms such as excitement versus gratitude.

Using the six emotion classification, we have adopted a Fleiss' Kappa statistical measure to assess the agreement of our ratings between the five members of this group. This statistical measure is perfect for assessing the reliability of agreement between our group of five raters, with categorical emotional labels. We created excel files for each member of our group to label and classify the chorus' lyrics into one of the six emotion categories. For any of the lyrics which were translated from non-English lyrics, there was also a note to indicate that the lyric was translated for the raters to take into consideration when trying to understand the lyrics. It is noted that the translation of the lyrics of non-English songs to English lyrics may present a lost in translation issue where an emotion is unable to be appropriately ascribed to the lyric. However, Genius' translated lyrics done up by the community attempts to reduce this error.

4. Finding the ground truth

To find our ground truth, we run our five files through the code we learnt from this module *'textprocessing.py'*. This helped aggregate the files into one excel file (*'agreement.xlsx'*) with the columns corresponding to each group member and the last (6th) column as the ground truth; and each row corresponding to the song and the rated emotion. The *'ground truth'* is then based on the most common emotion out of the five columns.

We acknowledge that this "majority rules" approach to finding the ground truth has limitations. Firstly, due to our small dataset of only five members, a ground truth is only found with at least three members labelling a certain emotion. After looking through our *agreement.xlsx* file, we realised our *ground truths* can swing drastically from one emotion to another, due to our small number of raters. To illustrate this, one ground truth can be concluded from a three-person agreement on the same emotion, while the other two agree on another emotion – the difference in the number of people agreeing is not big. In hindsight, an improvement could be made by ranking the six emotions (1-6) for each lyric, so as to prevent such a "swing".

4.1. Calculating Fleiss' Kappa

A Python statistical model library *'statsmodels'* was used to find the Fleiss' Kappa value. This was also inspired from *textprocessing.py*.

5. Creating Data Visualisations

To create the visualisations, we used the Python library '*Seaborn*'.

5.1. Bar chart

For our comparative analysis between TikTok and Billboard music, we created a barplot (Figure 3) with the x-axis as the six emotions and y-axis as the number of times each emotion-ground truth has appeared. We decided that this was the best way to enable easy comparison as the most frequent emotion within TikTok music itself, and also between TikTok and Billboard, can be easily told from a glance.

To create the bar chart, we first obtained the 'ground truths' for both TikTok and Billboard from the created agreement files, which were then placed inside a dataframe using Pandas. After obtaining the count for each emotion for TikTok and Billboard, we then used *Seaborn's barplot()* method to create the bar chart.

5.2. Confusion matrix

Adapting from previous works, we then run our chorus lyrics through an NLP model. This model was obtained from *pysentimiento*. We modified *textprocessing.py*, such that we could perform emotional analysis and obtain the predicted emotion based on what the model derives from the lyrics. We could then create confusion matrices (Figures 4 and 5) with the predicted emotion compared to the emotions in our ground truth similar to *textprocessing.py*.

However, we ran into a problem where the code we were dependent on required all six emotions to be in our ground truths but we did not necessarily have lyrics that might have evoked a certain emotion. To fix this issue, instead of changing our "ground truth" to include one from each label, we decided that adding a dummy row would be less intrusive to the original data. This might have been a problem due to our small dataset of less than 50 songs for each list, limited to the project's scope.

Analysis

Both TikTok and Billboard top 50 lists share about 20% (10 tracks) of the same songs. This follows our earlier understanding that viral songs on TikTok have a correlation to their charting positions as popular music.

1. Analysing Inter-rater Reliability

To measure our inter-rater reliability, we have decided to compute the Fleiss' Kappa value. The Fleiss' Kappa value ranges from 0 (zero agreement) to 1 (perfect agreement).

```
Agreement file saved at tt_output/agreement.xlsx
Fliess Kappa is 0.2546224961479199
```

Fig 1. Fleiss' Kappa Value for TikTok top 50 list

Our Fleiss' Kappa value for our Tiktok top 50 list (Figure 1) is only 0.255 (3 s.f.) which indicates a relatively low agreement.

```
Agreement file saved at bb_output/agreement.xlsx
Fliess Kappa is 0.3289583140770867
```

Fig 2. Fleiss' Kappa Value for Billboard top 50 list

Whereas, our Fleiss' Kappa value for our Billboard top 50 list (Figure 2) is 0.329 (3 s.f.) which is also relatively low agreement but higher than the former.

This low inter-rater reliability for both top 50 lists indicates that our ground truth is not considered reliable. We recognise that this is brought about due to the limitations of emotion ratings where human subjectivity is present – where one might find happiness in the lyrics, another may identify another emotion (Napier & Shamir, 2018).

2. Comparative Analysis of Emotions

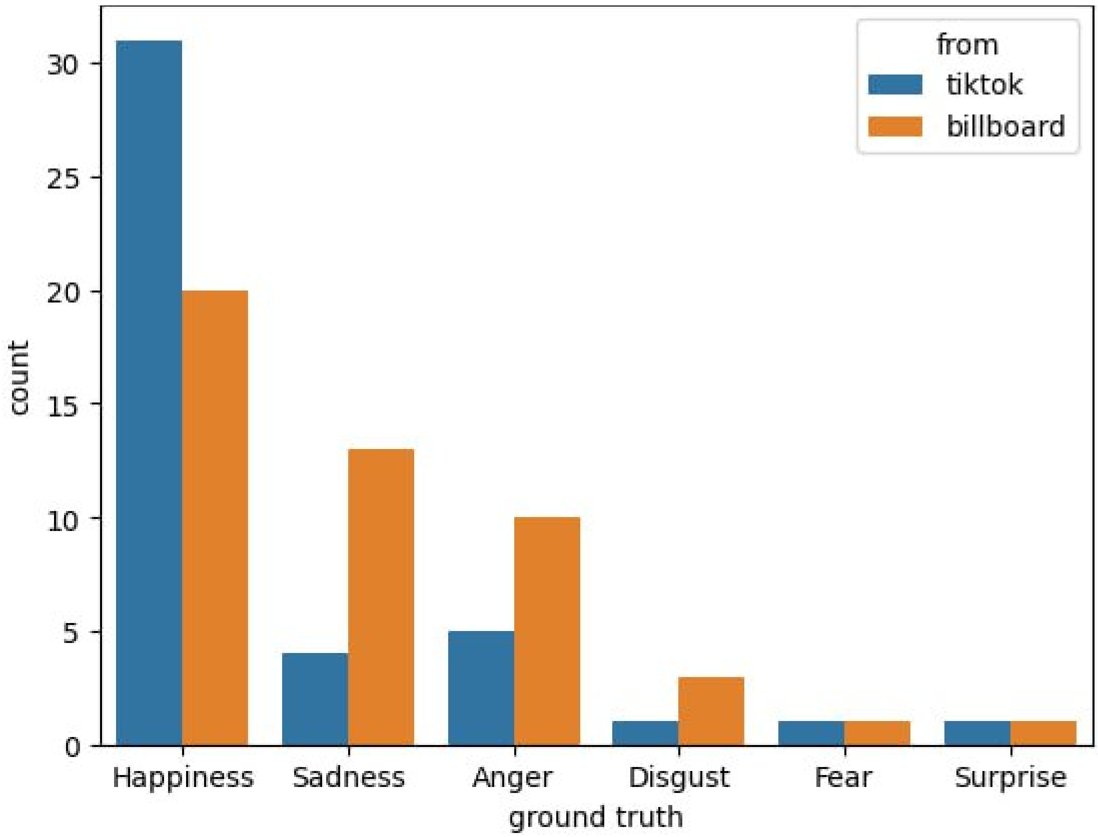

Fig 3. Bar chart measuring the six emotions versus the count of emotion-ground truth

This bar chart helps to directly compare the emotion of songs between TikTok and Billboard music. Happiness rated songs are highest for both lists. However, while "Happy" songs are the majority across all other emotions in TikTok music, this is contrasted by Billboard's which has a more diverse spread of songs in "Happiness", "Sadness" and "Anger". We can conclude that viral TikTok songs used in short-form videos tend to be more "Happy".

However, due to the limitation previously discussed (Methodology Section 5.2), the ground truth for Billboard and TikTok should have one less fear and one less surprise respectively as these were part of the dummy data used such that code could run with no errors.

3. Predictive Analysis

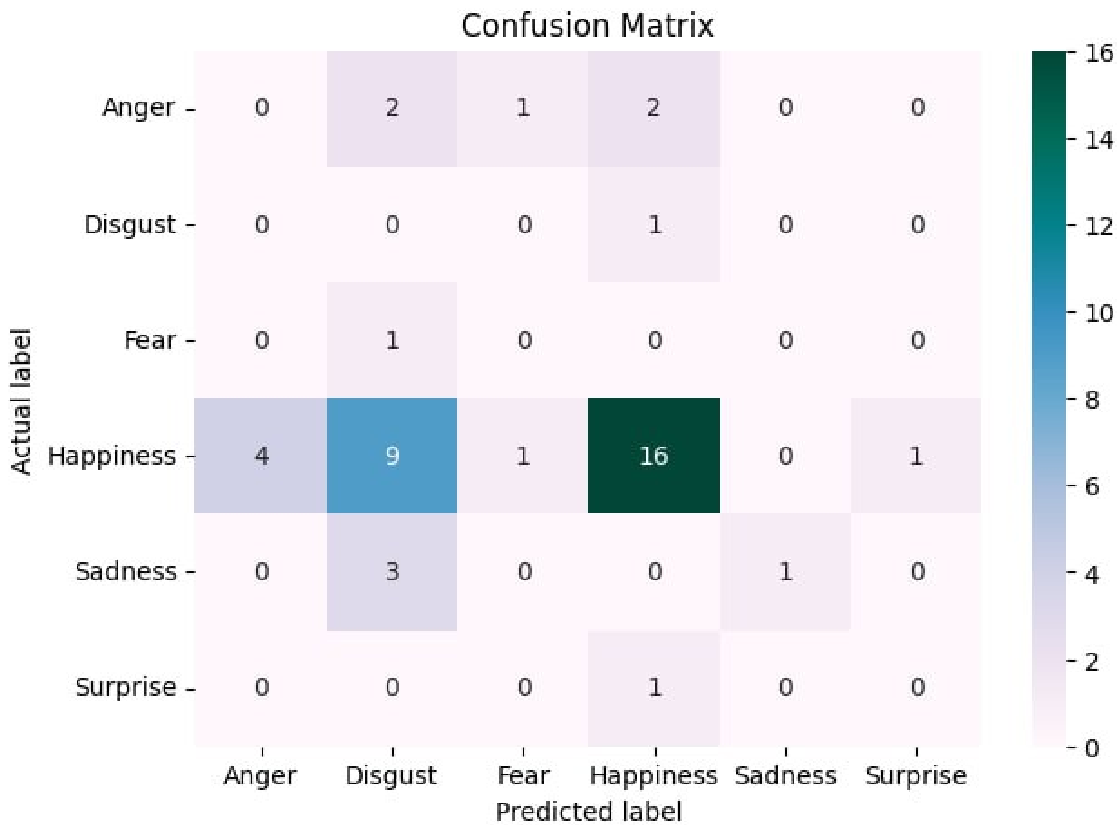

Fig 4. Confusion matrix for TikTok top 50 music

Now we examine how the existing NLP model from *pysentimiento* performs on the same dataset of lyrics. This aims to give a sense of how accurate the model is by comparing it with our group's ground truth. To do this, we created a multi-class confusion matrix.

As seen (Figure 4), the model is somewhat accurate when detecting happy lyrics as seen with the highest score of 16. However, even so, there exists a substantial discrepancy as seen from the

‘Happiness’ row where for actual Happiness labels, the predicted label may classify as Anger or Disgust instead. Additionally, for the songs with actual labels ‘Happiness’, the model has misclassified lyrics 15 times, which is about half the number of classified “Happy” songs (31). The model has seemingly misclassified lyrics of almost all other emotions.

Fig 5. Confusion matrix for Billboard Hot 100 music

Similarly for the Billboard Hot 100 music (Figure 5), the model is still the most accurate when detecting “Happy” lyrics. Most of the other emotions have been misclassified as well.

The model seems to have a bias for predicting lyrics that have been classified under “Disgust”. This could be seen (Figures 4 and 5) where the model predicts “Disgust” as the second most popular emotion, which is distinctly different from the number of actual “Disgust” labels. For example, while Billboard songs had 20 songs that were predicted to be "Disgust", only two were accurately predicted, leading to a 10% accuracy rate.

Conclusion

This report aims to take a more data-driven approach of comparing the emotions of lyrics in TikTok and popular music. Our approach differs by taking already trending songs on TikTok and popular music charts, before conducting an emotional analysis to discover any correlation. This approach reveals interesting conclusions of how popular songs, TikTok or not, gravitate towards having a "Happy" emotion. Furthermore, through the use of statistical models, we reveal the subjectivity of lyrics and how commercially trained NLP models for emotion analysis significantly deviates from our own ground truths.

(3000 words)

References

Adu, P. (2022). *Understanding the Use, Strengths and Limitations of Automated Text Analysis.* The Qualitative Report 2020. DOI:10.13140/RG.2.2.14156.16002

Balen, J. V., Burgoyne, J. A., Wiering, F., & Veltkamp, R. C. (n.d.). *AN ANALYSIS OF CHORUS FEATURES IN POPULAR SONG*. ISMIR. Retrieved November 24, 2023, from https://archives.ismir.net/ismir2013/paper/000180.pdf

Billboard. (n.d.). Billboard Charts Legend – Billboard. Billboard. Retrieved November 23, 2023, from https://www.billboard.com/billboard-charts-legend/

Billboard. (n.d.). TikTok Billboard Top 50 – Billboard. Billboard. Retrieved November 24, 2023, from https://www.billboard.com/charts/tiktok-billboard-top-50/

Billboard. (n.d.). Billboard Hot 100™ – Billboard. Billboard. Retrieved November 24, 2023, from https://www.billboard.com/charts/hot-100/

Coulter, A. (2022). Marketing Agile Artists: How Music Labels Can Leverage TikTok's Virality. https://doi.org/10.25101/22.5

Eichler, R. (n.d.). *TikTok: What It Is, How It Works, and Why It's Popular*. Investopedia. Retrieved November 23, 2023, from https://www.investopedia.com/what-is-tiktok-4588933

Ekman, P. (1992). *An Argument for Basic Emotions.* Lawrence Erlbaum Associates Limited. Retrieved November 24, 2023, from https://www.paulekman.com/wp-content/uploads/2013/07/An-Argument-For-Basic-Emotions.pdf

Genius. (2016). *How Genius Works*. https://genius.com/Genius-how-genius-works-annotated

Genius. (n.d.). *Community Policy & Moderation Guidelines*. https://genius.com/Genius-community-policy-and-moderation-guidelines-annotated

Genius. (2017). *How to Translate Songs on Genius*. https://genius.com/Genius-how-to-translate-songs-on-genius-annotated

Jamdar, A., Abraham, J., Khanna, K. & Dubey R. (2015). *Emotion Analysis of Songs Based on Lyrical and Audio Features.* International Journal of Artificial Intelligence & Applications (IJAIA). https://doi.org/10.48550/arXiv.1506.05012

Juslin, P. N. & Laukka, P. (2004). *Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening.* Journal of New Music Research 33:217238. DOI 10.1080/0929821042000317813.

Khan, M. (2021, January 21). How to Leverage Spotify API + Genius Lyrics for Data Science Tasks in Python. Medium. Retrieved November 24, 2023, from https://medium.com/swlh/how-to-leverage-spotify-api-genius-lyrics-for-data-science-tasks-in-python-c36cdfb55cf3

Lim, G. (2023, March 29). daily_dialog · What is emotion = 6? Hugging Face. Retrieved November 24, 2023, from https://huggingface.co/datasets/daily_dialog/discussions/3

Napier, K., & Shamir, L. (2018). Quantitative sentiment analysis of lyrics in popular music. *Journal of Popular Music Studies, 30*(4), 161-176. DOI:10.1525/jpms.2018.300411

Niblett, S. (2023, June 13). 'TikTokification': who wins when a song goes viral? Varsity. Retrieved November 23, 2023, from https://www.varsity.co.uk/music/25775

Saquete, E., Zubcoff, J., Gutiérrez, Y., Martínez-Barco, P., & Fernández, J. (2022). *Why are some social-media contents more popular than others? Opinion and association rules mining applied to virality patterns discovery.* Expert Systems with Applications, 197. *I*SSN 0957-4174. https://doi.org/10.1016/j.eswa.2022.116676

Schafer, K., & Eerola, T. (2020). *How listening to music and engagement with other media provide a sense of belonging_ An exploratory study of social surrogacy*. Psychology of Music, 48(2), 232-251. http://doi.org/10.1177/0305735618795036

Xu, L., Sun, Z., Wen, X., Huang, Z., Chao, C., Xu, L. (2021). *Using machine learning analysis to interpret the relationship between music emotion and lyric features. P*eer Journal of Computer Science. 7:e785. http://doi.org/10.7717/peerj-cs.785