# Literature Review

Title: Heuristic Approach For Generic Audio Data Segmentation And Annotation

Authors: Tong Zhang, C.-C. Jay Kuo

Publication Date: 1999

In this study, (Zhang et. al., 1999) examined an approach to audio classification. Their approach used a audio features and a rules-based approach to classify audio signals into speech, music, silence, etc.

Title: Audio Segmentation By Feature-Space Clustering Using Linear Discriminant Analysis And Dynamic Programming

Authors: M. Goodwin, J. Laroche

Publication: 2003 Ieee Workshop On Applications Of Signal Processing To Audio And Acoustics (Ieee Cat. No.03Th8684)

Publication Date: 2003

(Goodwin et. al., 2003) use linear discriminant analysis to cluster audio features. This method was used to separate audio signals into meaningful regions. Linear discriminant analysis was found to outperform principal component analysis for their applications.

Title: Algorithmic Clustering Of Music

Authors: Rudi L. Cilibrasi, P. Vitányi, R. D. Wolf

Publication: Proceedings Of The Fourth International Conference Onweb Delivering Of Music, 2004. Edelmusic 2004.

Publication Date: 2004

(Cilibrasi et. al., 2004) were able to use distinguish between musical genres and cluster pieces by composer using compression of strings that represent music. Their method used note-on and note-off event data imported from MIDI files. Data were then compressed and analysed for similarity using various distance measures. The clustered data was represented by converting the calculated distances into tree representations.

Title: A Music Recommender Based On Audio Features

Authors: Qing Li, Byeong Man Kim, Dong Hai Guan, Duk whan Oh

Publication: Sigir

Publication Date: 2004

(Li et. al., 2004) design a collaborative music recommended system which incorporates music audio features and metadata. Their method uses a clustering technique to integrate audio features into a “collaborative filtering framework” in order to improve recommendation performance.

Title: Content-Based Music Audio Recommendation

Authors: P. Cano, M. Koppenberger, N. Wack

Publication: Multimedia '05

Publication Date: 2005

(Cano et. al., 2005) introduce a metadata free system for recommending music using large datasets. Their method is based on a high-level music similarity metric, which accounts for tempo, meter, rhythm patterns, tonal aspects, timbre, and dynamics.

Title: Music Genre Classification Using MIDI And Audio Features

Authors: Zehra Cataltepe, Yusuf Yaslan, Abdullah Sonmez

Publication: Eurasip Journal On Advances In Signal Processing

Publication Date: 2007

(Cataltepe et. al., 2007a) used linear discriminant analysis and k-nearest neighbours to classify musical genres. Their analysis used MIDI files and audio features from MIDI, separately and combined. Audio features considered in this analysis included timbre, rhythm, and pitch content features.

Title: Music Recommendation Based On Adaptive Feature And User Grouping

Authors: Z. Cataltepe, B. Altinel

Publication: 2007 22Nd International Symposium On Computer And Information Sciences

Publication Date: 2007

In this study, (Cataltepe et. al., 2007b) introduced a framework for using different sets of audio features for each user to improve content recommendations. Using an entropy criterion, features where chosen for individual users. They also introduced an algorithm to learn the feature weights to give content based on user history.

Title: Predicting Music Popularity On Streaming Platforms

Authors: Carlos V. S. Araujo, Marco A. P. Cristo, Rafael Giusti

Publication: Anais Do Simpósio Brasileiro De Computação Musical (Sbcm 2019)

Publication Date: 2007

(Holt, 2007) predicted the popularity of music on streaming platforms. This analysis considers long-term popularity and virality. The findings are based on data from Spotify. Classification is performed using a Support Vector Machine model. Audio information was also considered. Audio features were calculated using the librosa package for Python on a 30 second sample of each song. Audio features considered include mel-frequency cepstral coefficients, spectral centroid, spectral flatness, zero crossings, and tempo. Findings show that popularity information alone is predictive of both long-term popularity and virality.

Title: A Probabilistic Music Recommender Considering User Opinions And Audio Features

Authors: Qing Li, Sung-Hyon Myaeng, Byeong Man Kim

Publication: Inf. Process. Manag.

Publication Date: 2007

(Li et. al., 2007) describe a music recommender system using probabilistic distributions and audio features. This system helps to alleviate user-bias, non-association, and cold start problems associated with sparse data. This study also showed how to combine primitive audio features into aggregate features using K-means and Fuzzy K-means clustering. Audio features considered include spectral centroid, spectral rolloff point, spectral flux, sum of scale factor, mel-frequency cepstral coefficients, rhythmic content features, and pitch content features.

Title: Novel Techniques For Audio Music Classification And Search

Authors: Kris West

Publication: Acm Sigmultimedia Records

Publication Date: 2008

(West, 2008) presents a methods to classify music by genre or “search-by-example.” In this doctoral thesis, West introduces procedures for parameterisation of music audio. Additionally, the thesis introduces a novel machine learning algorithm based on a Decision Tree algorithm to process multi-variate audio features. Finally, a number of methods are introduces for extending music classifiers to estimate musical similarity. Methods introduced in this thesis allow for more efficient music search by reducing computational complexity and improving performance.

Title: Exploring Relationships Between Audio Features And Emotion In Music

Authors: C. Laurier, O. Lartillot, T. Eerola, P. Toiviainen

Publication Date: 2009

(Laurier et. al., 2009) examine the association between emotion categories and audio features. The analysis uses Support Vector Regression to estimate the relationship between audio descriptors and emotional categories as determined in psychological studies. Audio features considered in this analysis include timbral, tonal, and rhythmic features comprised of 200 features statistics.

Title: Emotion-Based Music Retrieval On A Well-Reduced Audio Feature Space

Authors: Maria M. Ruxanda, Bee Yong Chua, Alexandros Nanopoulos, Christian S. Jensen

Publication: 2009 Ieee International Conference On Acoustics, Speech And Signal Processing

Publication Date: 2009

(Ruxanda et. al., 2009) examine dimensionality reduction as a means to increase performance for real-time music retrieval. The approach projects the music into a music emotion feature space, which is dimensionally reduced, improving retrieval performance while preserving accuracy. A variety of dimensionality reduction techniques were evaluated.

Title: Clustering And Classification Of Music By Interval Categories

Authors: Aline K. Honingh, Rens Bod

Publication: Mcm

Publication Date: 2011

(Honingh et. al., 2011) clustered and classified music using six different musical characters: semitones, whole-tones, diminished triads, augmented triads, diatonic scale, and tritones. The technique was used for composer clustering and genre classification

Title: Ameliorating Music Recommendation: Integrating Music Content, Music Context, And User Context For Improved Music Retrieval And Recommendation

Authors: Markus Schedl

Publication Date: 2013

(Schedl, 2013) combined music content with user-centric information to improve music retrieval and recommendations. Notably, this analysis incorporates geographical places of interest into the retrieval process.

Title: Comparing Audio Features And Playlist Statistics For Music Classification

Authors: Igor Vatolkin, Geoffray Bonnin, Dietmar Jannach

Publication: Ecda

Publication Date: 2014

(Vatolkin et. al., 2014) consider feature extraction from user-shared playlists as an alternative to computationally costly audio feature extraction. The study outlines strategies to classify genre and style using large collections of user-provided playlists.

Title: End-To-End Learning For Music Audio

Authors: Sander Dieleman, Benjamin Schrauwen

Publication: 2014 Ieee International Conference On Acoustics, Speech And Signal Processing (Icassp)

Publication Date: 2014

(Dieleman et. al., 2014) attempt to apply feature learning techniques directly to raw audio signals, in contrast to mid-level features such as spectrograms typically used in convolutional neural networks. The performance of this approach did not match the spectrogram-based approach, but the techniques did work on raw audio.

Title: Content-Based Music Recommendation Using Underlying Music Preference Structure

Authors: M. Soleymani, Anna Aljanaki, F. Wiering, R. Veltkamp

Publication: 2015 Ieee International Conference On Multimedia And Expo (Icme)

Publication Date: 2015

(Soleymani et. al., 2015) propose a solution to the cold start problem, wherein new users or items are a challenge to recommender systems. The method classifies and recommends songs based on a five-factor preference vector estimated using audio features. The preference vector includes Mellow, Unpretentious, Sophisticated, Intense and Contemporary (MUSIC). The method outperforms genre-based and user-based recommendations while decreasing popularity-bias in recommended songs.

Title: Revisiting The Problem Of Audio-Based Hit Song Prediction Using Convolutional Neural Networks

Authors: Li-Chia Yang, Szu-Yu Chou, Jen-Yu Liu, Yi-Hsuan Yang, Yian Chen

Publication: 2017 Ieee International Conference On Acoustics, Speech And Signal Processing (Icassp)

Publication Date: 2017

(Yang et. al., 2017) attempt to jointly learning audio features and predicting popularity using a convolutional neural network. The convolutional neural network used low-level mel-spectrogram directly as input for feature learning, as well as an external audio-tagging dataset. This work suggest that deep structures are more accurate than shallow structures in predicting the popularity of music.

Title: Music Popularity: Metrics, Characteristics, And Audio-Based Prediction

Authors: Junghyuk Lee, Jong-Seok Lee

Publication: Ieee Transactions On Multimedia

Publication Date: 2018

(Lee et. al., 2018) define eight popularity metrics covering multiple aspects of popularity (Debut, Max, Mean, Std, Length, Sum, Skewness, and Kurtosis), and then analyse each metric with long-term real-world chart data. The study examined popularity metrics for the Billboard Hot 100 chart between 1970 and 2014. The study combined music complexity features with conventional audio features (such as MPEG-7 and Mel-frequency cepstral coefficient) to build a Support Vector Machine classification model. Results demonstrate that popularity can be predicted based on audio features alone.

Title: Determining Characteristics Of Popular Local Songs In Indonesia'S Music Market

Authors: Limisgy Ramadhina Febirautami, Isti Surjandari, Enrico Laoh

Publication: 2018 5Th International Conference On Information Science And Control Engineering (Icisce)

Publication Date: 2018

(Febirautami et. al., 2018) use the Spotify API to obtain audio features (Acousticness, Danceability, Energy, Instrumentalness, Key, Liveness, Mode, Speechiness, Tempo, Time signature, and Valence). Audio features analysed using Decision Trees to identify attributes representing popular songs.

Title: Predicting Hit Songs with Machine Learning

Authors: Minna Reiman; Philippa Örnell

Publication: EXAMENSARBETE INOM TEKNIK, GRUNDNIVÅ, 15 HP

Publication Date: 2018

(Reiman et.al., 2018) predict popularity of hit songs using audio features from the Spotify API as inputs to Logistic Regression, K-Nearest Neighbours, Gaussian Naïve Bayes, and Support Vector Machine models. The study was not able to accurately predict song popularity, potentially due to an overly diverse dataset for non-hit songs. In this study popularity was defined as an appearance on the Billboard Hot 100 charts.

Title: Contextual Personalized Re-Ranking Of Music Recommendations Through Audio Features

Authors: Boning Gong, Mesut Kaya, Nava Tintarev

Publication: Arxiv

Publication Date: 2020

(Gong et. al., 2020) propose a contextual re-ranking algorithm based on audio features. The study compares global and personalized models. User-specific models were found to outperform the global model. Notably, the study found a correlation between audio features and contextual conditions, such as time of day.

Title: A Multimodal End-To-End Deep Learning Architecture For Music Popularity Prediction

Authors: David Martín-Gutiérrez, Gustavo Hernández Peñaloza, Alberto Belmonte-Hernández, Federico Álvarez García

Publication: Ieee Access

Publication Date: 2020

(Martín-Gutiérrez et. al., 2020) develop a dataset to aid researchers in correlating different datasets, which use different metrics to evaluate popularity. Additionally, a Deep Learning architecture is used to predict popularity based on audio features, lyrics, and metadata.

Title: Effect Of Feature Selection On The Accuracy Of Music Genre Classification Using Svm Classifier

Authors: De Rosal Ignatius Moses Setiadi, Dewangga Satriya Rahardwika, Eko Hari Rachmawanto, Christy Atika Sari, Ajib Susanto, Ibnu Utomo Wahyu Mulyono, Erna Zuni Astuti, Amiq Fahmi

Publication: 2020 International Seminar On Application For Technology Of Information And Communication (Isemantic)

Publication Date: 2020

(Setiadi et. al., 2020) analyse the effects of features selection on the accuracy of music genre classification. Support Vector Machines are used to classify genres based on combination sets of audio features from the Spotify API. Classification of genre was not found to improve in accuracy by removing features, but the top 13 features have similar accuracy to all 17 features, potentially reducing calculation requirement when classifying future genres. The features which were not found to be predictive of genre are Time signature, Track name, Key, and Track id.

Title: Music Personalized Label Clustering And Recommendation Visualization

Authors: Yongkang Huo

Publication: Complex.

Publication Date: 2021

(Huo, 2021) introduce a tag-based collaborative filtering algorithm optimised to improve the accuracy of recommendations.

Title: Music Popularity Prediction Through Data Analysis Of Music’S Characteristics

Authors: Jae-Hyuck Kim

Publication: International Journal Of Science, Technology And Society

Publication Date: 2021

(Kim, 2021) examines music popularity on Spotify using data from 2010 to 2019. This paper computes the averages of audio features to investigate rankings on Spotify. Audio features investigated in this study include BPM, Energy, Danceability, Loudness, Liveness, Valence, Length, Acousticness, and Speechiness. Correlation between audio features and popularity were investigated. Linear Regression, K-Nearest Neighbours, and Random Forest models were used to predict popularity based on audio features and genre.

Title: A Clustering Analysis Method For Massive Music Data

Authors: Yanping Xu, Sen Xu

Publication Date: 2021

(Xu et. al., 2021) introduce a method to cluster music data using the K-means Clustering on extracted spectral data extracted using a Fast Fourier Transform.

Title: Music Mood Classification System For Streaming Platform Analysis Via Deep Learning Based Feature Extraction

Authors: Yu-Chia Chen, Zih-Ching Chen, Chih-Hsien Hsia

Publication: 2021 Ieee International Conference On Consumer Electronics-Taiwan (Icce-Tw)

Publication Date: 2021

(Chen et. al., 2021) propose a deep learning model for extracting features from audio and lyrics in order to classify songs by mood. They inputted audio features provided in the Spotify API and lyrics crawled from Genius for music mood classification. They found that their proposed deep learning classification method outperformed a number of conventional machine learning approached (Logistic Regression, XGBoost, Decision Tree Classifiers, Random Forest Classifiers, Bert for Sequence Classification).

Title: Music-Circles: Can Music Be Represented With Numbers?

Authors: Seokgi Kim, Jihye Park, Kihong Seong, Namwoo Cho, Junho Min, Hwajung Hong

Publication: Arxiv

Publication Date: 2021

(Kim et. al., 2021) created a data visualisation tool using music clustering based on audio features provided in the Spotify API. The clustering used in this analysis was K-means Clustering.

Title: Learning Recommendation Algorithm Based On Improved Bp Neural Network In Music Marketing Strategy

Authors: Lei Li

Publication: Computational Intelligence And Neuroscience

Publication Date: 2021

(Li, 2021) propose a method for learning user music preferences in order to improve recommendation algorithms. Convolutional neural networks to extract high-level features for songs which are used as input to a probabilistic graph model.

Title: Analysis Of Sonic Effects Of Music From A Comprehensive Datasets On Audio Features

Authors: Tobechukwu Okechukwu Otuokere, Agbotiname Lucky Imoize, Aderemi Aaron-Anthony Atayero

Publication Date: 2021

(Otuokere et. al., 2021) statistically analyse 160,000 songs from 1921-2020 using audio features provided by the Spotify API. The analysis explore descriptive statistics for audio features, correlation between audio features, variation of audio features by year, and song data for top artists for multiple cohorts.

Title: Catching the Earworm: Understanding Streaming Music Popularity Using Machine Learning Models

Authors: Andrea Gao

Publication: E3S Web of Conferences 253, 03024

Publication Date: 2021

(Gao, 2021) analyses popularity in streaming music using machine learning models. The analysis uses Linear Regression, Logistic Regression, Decision Tress, Random Forests, Boosting Trees with Principal Component Analysis, and Neural Networks to predict popularity using audio features from the Spotify API. The analysis uses 130,663 tracks from Spotify collected in 2018 and 2019.

Title: Music Recommendation Algorithm Based On Multidimensional Time-Series Model Analysis

Authors: Juanjuan Shi

Publication: Complex.

Publication Date: 2021

(Shi, 2021) introduce a music recommendation algorithm approach to consider users’ long-term, medium-term, and short-term behaviours. The algorithm represents the users’ behaviour as a multidimensional timer series, and uses a recurrent neural network to provide music recommendations.

Title: Statistical And Visual Analysis Of Audio, Text, And Image Features For Multi-Modal Music Genre Recognition

Authors: Ben Wilkes, Igor Vatolkin, Heinrich Müller

Publication: Entropy (Basel, Switzerland)

Publication Date: 2021

(Wilkes et. al., 2021) classify musical genre using multiple modalities, namely audio, text, images of audio signals, cover art, and lyrics. Different combinations of engineered features are used as input to Naïve Bayes, Support Vector Machine, and Random Forest classification models. Statistical analysis of results indicate that incorporating multiple modalities leads to better classification of results.

Title: A Music Emotion Classification Model Based On The Improved Convolutional Neural Network

Authors: Xiaosong Jia

Publication: Computational Intelligence And Neuroscience

Publication Date: 2022

(Jia, 2022) use a bi-directional long short-term memory network, a convolutional recurrent neural network, and a softmax classification function to classify music by emotion. The model combines low-level audio features mel-frequency cepstral coefficient and residual phase with audio spectrograms.

Title: Music Classification Method Using Big Data Feature Extraction And Neural Networks

Authors: Xiabin Li, Jin Li

Publication: Journal Of Environmental And Public Health

Publication Date: 2022

(Li et. al., 2022) use a convolutional neural network to create audio features for music which can be used in recommendation algorithms.

Title: Novelty And Cultural Evolution In Modern Popular Music

Authors: Katherine O'Toole, Em?ke-Ágnes Horvát

Publication: Arxiv

Publication Date: 2022

(O'Toole et. al., 2022) compare established characteristics for musical genres with identified novel artifacts in order to study the relationship between novelty and commercial success. The analysis combines Music Information Retrieval data with lyrics for Billboard Hot 100 songs to calculate a novelty score for each song. Novelty scores were created based on Euclidean and Mahalanobis distances between song and release year average genre values for lyrical feature and audio features, respectively. This study was limited to songs that appeared on the Billboard Hot 100.

## Notes

Although no sources have been found to examine the research questions for this report, clustering and popularity prediction have been successful in separate studies. Insights into clustering and prediction models appear to indicate that neural networks typically perform the best, but simpler methods also lead to meaningful results (SOURCES).

It is worth noting that when predicting popularity, some studies had increased success separating tracks by genre (SOURCES). This may be easier than the proposed method of clustering based on audio features.

Also worth noting, audio feature statistics were found by (SOURCE) to be relatively stable over time when controlling for genre. This could be a useful simplifying assumption to reduce computational requirements for this project. However, if time allows, clustering in terms of genre, year, and/or audio features could allow for more accurate and insightful predictions.