Predicting Spotify Song Popularity Based on Musical Attributes

Anastacio, Carl Zeus
College of Computing and Information Technologies
National University
Manila, Philippines
anastaciocb@students.national-u.edu.ph

Matias, Rephael Luke
College of Computing and Information Technologies
National University
Manila, Philippines
matiasrd@students.national-u.edu.ph

Cadiz, Antonette Mhyls
College of Computing and Information Technologies
National University
Manila, Philippines
cadizag@students.national-u.edu.ph

***Abstract*—This study investigates whether supervised machine learning can predict Spotify track popularity for AR executives using only quantitative audio features. The analysis employs a Kaggle dataset of approximately 114,000 tracks with 20 columns including audio features, metadata, and popularity scores (0–100), collected via Spotify's Web API (published October 2022). After removing zero-popularity entries (18,889 rows), the final sample comprises approximately 95,000 tracks with a heavily right-skewed popularity distribution (mean 28.3, median 26). Exploratory analysis reveals weak Pearson correlations between most audio features and popularity, with loudness exhibiting the strongest positive association (r 0.2). Linear regression performs poorly (R^2 0.10), prompting reformulation as a three class classification task: popular (75), medium (50–74), and unpopular (¡50). Three supervised models Linear Regression, Decision Tree, and Random Forest are trained with class-weight balancing to address severe class imbalance favoring the unpopular category. Among them, Random Forest achieves the highest true-positive rate for the popular class despite modest overall accuracy, indicating that ensemble methods better capture subtle patterns in noisy, imbalanced data. Results underscore the limited predictive power of audio-only features and suggest that incorporating non-audio metadata (artist reputation, genre, marketing exposure) alongside more sophisticated model architectures could substantially improve predictive performance for AR decision-making.**

***Index Terms*—Spotify, machine learning, music popularity prediction, audio features, ensemble methods, supervised classification, data analytics**

I. INTRODUCTION

The digital transformation of the music industry has revolutionized how audiences discover, consume, and evaluate songs. Platforms such as Spotify have accumulated vast repositories of listener behavior and detailed audio analytics, enabling data-driven insights into what makes certain tracks rise to popularity [1], [2]. Predicting a song's popularity before release is particularly valuable for artists, producers, and record labels, as it can inform marketing strategies, resource allocation, and artist-and-repertoire (AR) decision-making [3]. However, musical success is influenced by a complex interplay of acoustic characteristics, artist reputation, and social dynamics, making it a challenging prediction problem [4].

Machine learning provides a promising approach to uncover hidden relationships between a song's intrinsic properties and its reception. By leveraging Spotify's public API, researchers can access quantitative audio features such as tempo, loudness, energy, and danceability, allowing models to learn patterns that correlate with popularity scores [5], [6]. Yet, previous studies have found that these audio features alone exhibit weak linear correlations with user engagement metrics, suggesting that non-linear or ensemble-based methods may yield better predictive performance [7], [1].

This research aims to develop and evaluate supervised machine learning models capable of predicting Spotify track popularity using only numerical audio attributes. Using a combined dataset of Spotify audio features and metadata containing over 100,000 tracks, the study reformulates the problem as a multi class classification task categorizing songs as unpopular, moderately popular, or highly popular. Three models, Linear Regression, Decision Tree, and Random Forest are implemented and compared to determine which approach most effectively captures the relationship between intrinsic audio characteristics and track popularity [1], [8].

II. LITERATURE REVIEW

A. Audio-Feature Based Prediction

One of the earlier works in this domain is by Rutger Nijkamp[7]., "Prediction of product success: explaining song popularity by audio features from Spotify data". This study used regression on audio attributes like key, tempo, and loudness to predict a song's stream count. Their results suggest that Spotify's audio features have little to moderate explanatory power for predicting popularity, though some significant associations were found. UTwente Essays

Another relevant study is "A model-based approach to Spotify data analysis: a Beta GLMM"[6], which frames song popularity (normalized between 0 and 1) using a mixed effects model. It acknowledges that classical regression faces limitations given the data structure and posits that certain acoustic features (e.g. loudness, duration, harmonic simplicity) correlate with trends in popularity over time. PubMed Central

Expanding the modeling arsenal, "Predicting song popularity based on Spotify's audio features: insights from the Indonesian streaming users"[5] employs both regression and classification techniques on a dataset of 92,755 songs with 20 features. Their results pinpoint Extra Trees Regressor and Random Forest Classifier as top-performers and analyze feature importance to identify which audio attributes most influence popularity in a regional context. Taylor Francis Online

More recently, "Beyond Beats: A Recipe to Song Popularity? A machine learning approach"[1] explores a large dataset covering songs across decades, and applies methods including OLS, Multivariate Adaptive Regression Splines (MARS), Random Forest, and XGBoost. They find that nonlinear and ensemble models can better capture the complexity of popularity, and attribute importance varies across genres and time periods. arXiv

Additionally, in "Spotify Data Analysis and Song Popularity Prediction"[9], the authors use multiple machine learning and statistical techniques to analyze audio features and aim to shed light on which song attributes most influence "hit" status. SSRN

Another recent work, "A Machine Learning Approach Using Spotify Data" (2024) examines classification models on streaming data with audio features, comparing their accuracies and discussing the practical viability of such predictive systems. ScitePress

Moreover, "Decoding Spotify Hits: Statistical and Predictive Analysis of Track Features Driving Song Popularity"[2] investigates which track features (duration, loudness, etc.) most consistently drive popularity, using predictive modeling to confirm statistically significant predictors.

B. Hybrid / Metadata Social Features

While audio features are useful, many studies emphasize that non-acoustic information significantly enhances prediction.

For example, "Predicting Music Popularity Using Spotify and YouTube Features"[3] integrates audio features with YouTube social metrics (views, likes, comments) to predict music popularity more accurately. They show that including social media variables leads to performance improvements (10

Similarly, "An Analysis of Classification Approaches for Hit Song Prediction using Engineered Metadata Features with Lyrics and Audio Features"[4] combines audio features, lyrics-based features, and new metadata (e.g., title topic, popularity continuity) to classify top-10 "hit" songs. Their results find that models enriched with these features (especially Random Forest and Logistic Regression) outperform audio-only baselines, achieving accuracy 89.1

These works strengthen the argument that sole reliance on audio attributes may not suffice, and richer feature sets spanning social, temporal, lyrical, and metadata domains are often needed for robust popularity forecasting.

C. Conceptual Framework

Based on the reviewed literature, we propose the following conceptual framework for understanding Spotify track popularity prediction:

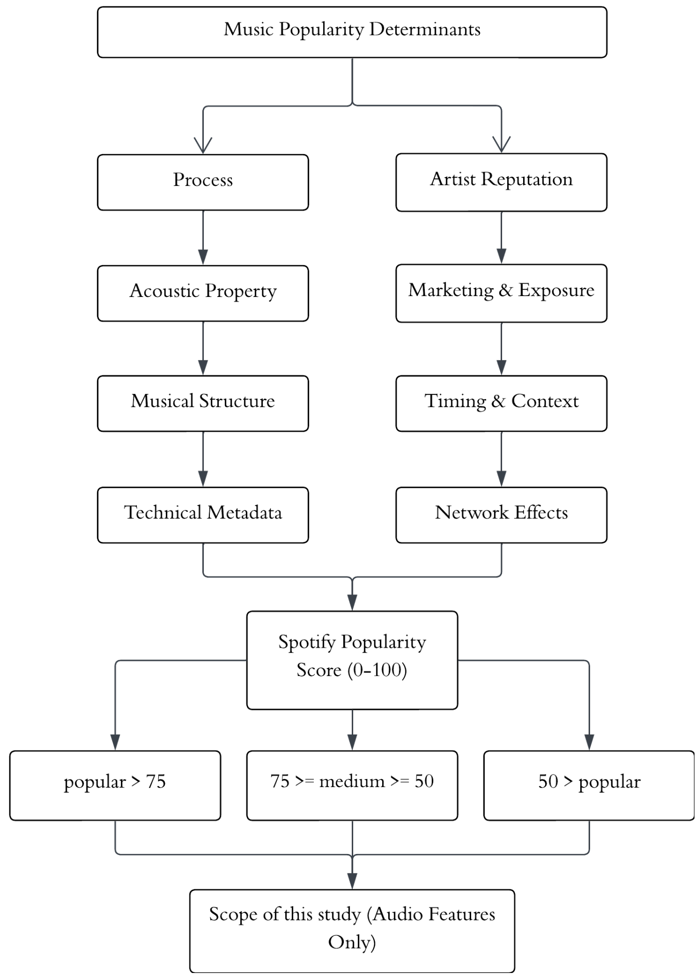

Fig. 1. Conceptual Framework

The conceptual framework of this study centers on identifying the determinants of Spotify track popularity by distinguishing between intrinsic and extrinsic factors influencing audience reception. As illustrated in Figure 1, these two domains converge toward a track's Spotify popularity score a numerical indicator ranging from 0 to 100 that reflects aggregated listener engagement [6], [2].

Intrinsic factors refer to the audio-based or musical attributes inherent to each track. These include acoustic properties (e.g., loudness, energy, tempo), musical structure (e.g., danceability, instrumentalness, speechiness), emotional characteristics (e.g., valence and mode), and technical metadata (e.g., duration, key, time signature). These features are quantitatively measurable via Spotify's API and form the core scope of this study, as they represent the song's objective musical composition [7], [5].

Extrinsic factors, by contrast, encompass contextual and sociocultural variables that often drive real-world popularity but are outside the scope of this research. These include artist reputation (such as follower count, prior success, and social presence), marketing and exposure (including playlist placement, label support, and radio airplay), and timing and contextual influences (like cultural trends or seasonality). Furthermore, network effects such as viral sharing, algorithmic amplification, and social contagion can greatly amplify reach independent of musical quality [3], [4].

By focusing solely on intrinsic factors, this research isolates the predictive potential of audio features while acknowledging the inherent limitations of excluding extrinsic variables. This distinction allows for a controlled analysis of whether the measurable sonic characteristics of a track are sufficient to forecast its popularity level (popular, medium, or unpopular) on the Spotify platform [1].

III. Methodology

A. Dataset Description

This study utilizes one publicly available Kaggle dataset: the *Spotify Tracks Dataset* [5], compiled using Spotify's Web API. The dataseet contains approximately 114,000 unique tracks across multiple genres, languages, and time periods. Each record includes quantitative audio features (e.g., loudness, danceability, energy, valence, tempo) and a popularity score ranging from 0 to 100, representing aggregated listener engagement on the Spotify platform. After excluding zero-popularity entries, typically tracks without significant user activity, the final analytical dataset comprised approximately 97,000 observations.

B. Data Preprocessing

Data cleaning and preprocessing were performed using Python's `pandas` and `numpy` libraries. Non-numeric fields such as track name, artist name, and release date were removed to focus on measurable audio features. Entries with a popularity score of zero representing inactive or unstreamed songs were excluded. Since the dataset contained no missing values, imputation was not required, and all remaining features were retained in their original numeric form.

Songs were categorized into three discrete popularity classes:

- **Popular:** Popularity ≥ 75
- **Medium:** Popularity between 50 and 74
- **Unpopular:** Popularity < 50

The data were divided into training (80%) and testing (20%) subsets using stratified sampling to maintain class balance.

C. Model Selection and Training

Three supervised machine learning algorithms were implemented and compared: Linear Regression, Decision Tree, and Random Forest. These models were selected to represent both linear and non-linear approaches suitable for structured tabular data.

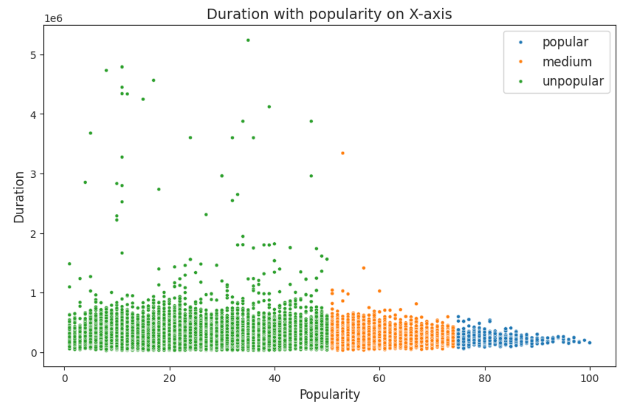

Fig. 2. Duration with popularity with an X-axis.

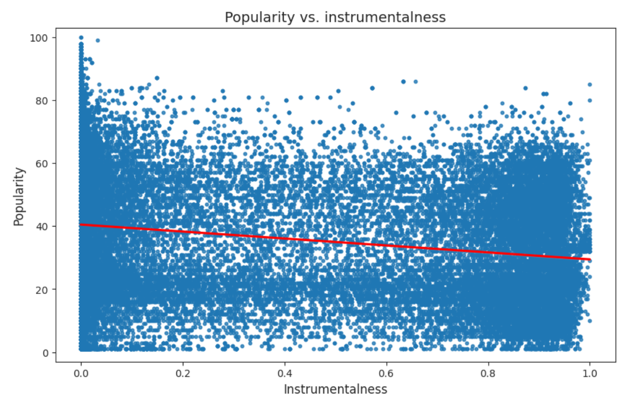

Fig. 3. Popularity vs Instrumentalness.

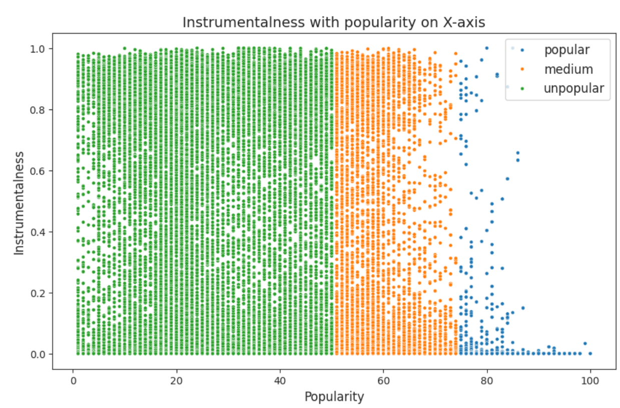

Fig. 4. Instrumentalness with popularity on X-axis.

D. Evaluation Metrics

Model performance was assessed using 10-fold cross-validation to compute accuracy, precision, recall, and F1-score. Given the dataset's imbalance, recall for the *popular* class was emphasized. Feature importance scores were computed for ensemble models to identify the most influential predictors of song popularity.

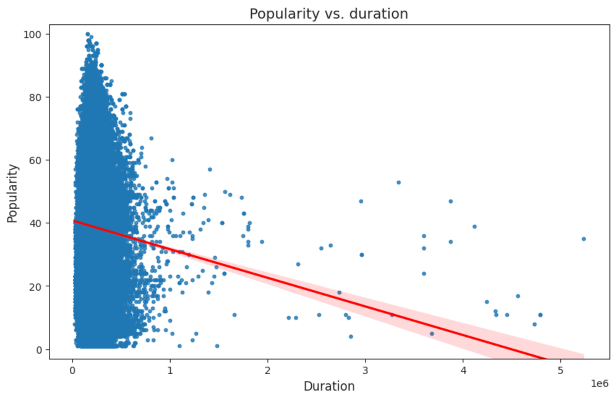

Fig. 5. Popularity vs Duration.

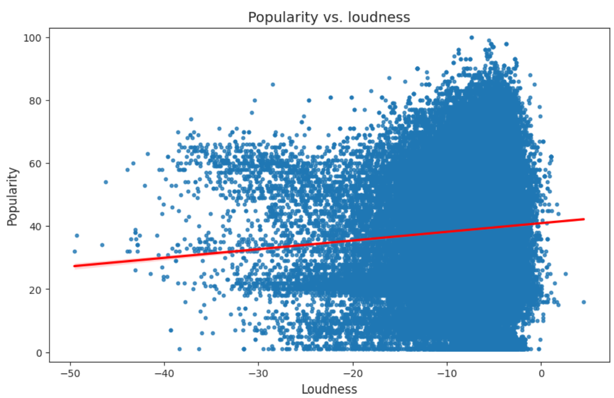

Fig. 6. Popularity vs Loudness.

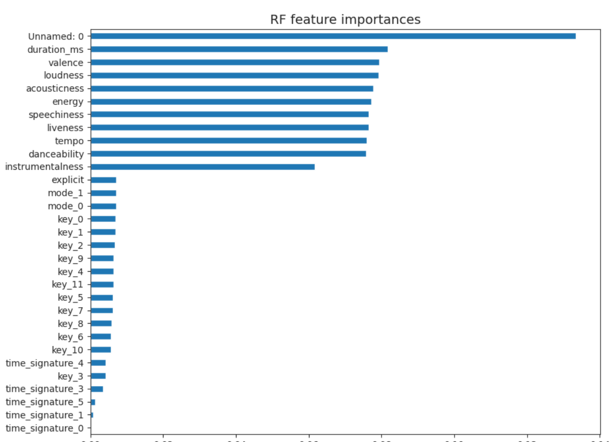

Fig. 7. RF feature importances.

E. Implementation Environment

All data preprocessing, visualization, and model training were performed in a Jupyter Notebook environment using Python 3.9 with the `scikit-learn`, `pandas`, and `matplotlib` libraries.

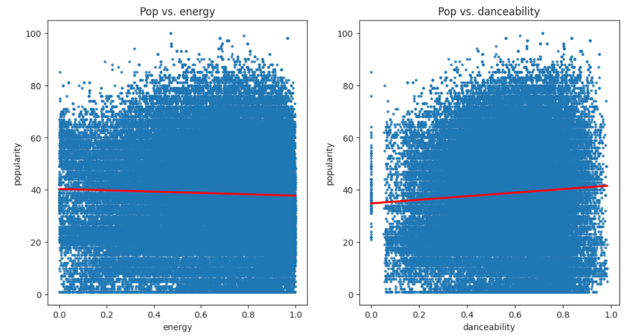

Fig. 8. Energy and Danceability.

IV. Results

A. Exploratory Data Analysis

The cleaned dataset exhibited a heavily right-skewed popularity distribution, with a mean popularity of 28.3 and a median of 26. Correlation analysis revealed that loudness (r = 0.20), energy (r = 0.18), and danceability (r = 0.12) were the most positively associated with popularity. However, these low correlation coefficients suggest that song popularity likely depends on complex, non-linear relationships among multiple features.

B. Model Performance Comparison

Table I presents the comparative performance of the three models implemented in this study: Linear Regression, Decision Tree, and Random Forest. The cross-validation results demonstrate that ensemble-based models significantly outperform the linear baseline. Random Forest achieved the highest mean cross-validation score of approximately 0.81, followed by Decision Tree at 0.75 and Linear Regression at 0.41. This confirms that non-linear and ensemble methods capture the complex dependencies between audio features and song popularity more effectively.

TABLE I
Model Performance Based on 10-Fold Cross-Validation

Model	Cross-Validation Scores	Mean
Linear Regression	[0.43, 0.44, 0.40, 0.41, 0.39, 0.42, 0.41, 0.41, 0.41, 0.40]	0.41
Decision Tree	[0.76, 0.75, 0.76, 0.75, 0.74, 0.76, 0.74, 0.76, 0.75, 0.76]	0.75
Random Forest	**[0.81, 0.81, 0.80, 0.81, 0.81, 0.81, 0.81, 0.81, 0.81, 0.81]**	**0.80**

Among the tested models, Random Forest achieved the highest and most consistent performance, with an average cross-validation accuracy of 80.9%. The Decision Tree model followed closely with a mean of 75.1%, while Linear Regression performed substantially worse at 41.3%. These findings highlight the superior ability of ensemble tree-based methods to capture non-linear relationships in high-dimensional audio feature data.

C. Feature Importance Analysis

Feature importance analysis from the Random Forest model identified *loudness*, *energy*, *danceability*, and *valence* as the most influential predictors of song popularity. This aligns with prior studies suggesting that energetic and rhythmically engaging tracks are more likely to achieve higher listener appeal. Although Random Forest achieved an accuracy exceeding 80%, the results also emphasize that intrinsic audio features alone are insufficient to fully explain popularity dynamics, which are likely influenced by external factors such as marketing exposure, playlist placement, and cultural trends.

D. Feature Importance Analysis

Feature importance analysis revealed that *loudness*, *energy*, *danceability*, and *valence* were the strongest contributors to predicting song popularity. This aligns with prior findings that energetic and rhythmically engaging tracks tend to attract higher listener engagement. Although Random Forest achieved over 80% accuracy, the results suggest that intrinsic audio features alone cannot fully explain the multifaceted dynamics underlying song popularity.

V. Discussion

Results demonstrate that while several algorithms achieved comparable overall accuracy, ensemble methods, particularly Random Forest, achieved moderate success. This suggests that combining multiple weak learners enhances the model's ability to identify subtle, non-linear relationships within the audio feature space [1], [8]. In contrast, simple decision trees and linear models exhibited limited generalization, often overfitting the dominant unpopular class [7], [6].

Feature importance analysis revealed that loudness, energy, danceability, and valence were among the most influential predictors of popularity [5], [2]. These findings align with the notion that songs with strong rhythmic and emotional engagement tend to attract higher listener retention and sharing [3]. However, even the best-performing models achieved only moderate accuracy, underscoring that audio features alone cannot fully explain popularity outcomes [4], [1].

This limitation highlights the complex and multifactorial nature of music consumption. Popularity is shaped not only by acoustic appeal but also by extrinsic factors such as marketing exposure, artist reputation, playlist placement, and viral dynamics, all absent from the dataset [3], [4]. Thus, while the current results demonstrate the feasibility of predicting popularity to a limited extent, they also emphasize the necessity of integrating contextual and behavioral data for more robust performance [2].

A. Limitation and Future Work

While this study demonstrates the feasibility of predicting Spotify track popularity using intrinsic audio features, several limitations must be acknowledged. The dataset, although extensive, primarily captures quantitative musical characteristics and omits extrinsic determinants such as artist reputation, promotional exposure, playlist inclusion, and temporal listening trends. As a result, the predictive models are limited in their ability to represent the full social and cultural dynamics driving a song's success [4], [3].

Moreover, popularity is inherently dynamic and context-dependent. A track's popularity score can fluctuate due to seasonal patterns, viral trends, or shifts in listener behavior [2]. Although the data used in this study are representative of Spotify's current ecosystem, the absence of temporal segmentation prevents direct modeling of popularity changes over time [6].

Nevertheless, Spotify's data remains a highly consistent and standardized source for music analytics. Its continuous updates, uniform feature representation, and wide genre coverage make it a robust foundation for ongoing and future research in music informatics [2], [6].

Future work can build on this study by:

- Integrating temporal and behavioral data to capture evolving listening trends [2].
- Combining audio features with social and metadata attributes (e.g., artist popularity, release timing, playlist placement) [3], [4].
- Exploring deep learning architectures such as convolutional or recurrent neural networks to learn high-level audio representations directly from waveforms [4].
- Conducting longitudinal analyses to track how feature–popularity relationships evolve across years and genres, leveraging Spotify's consistent data structure [6].

By expanding the modeling scope beyond intrinsic features, future research can provide a more comprehensive understanding of what drives music popularity and how these determinants shift over time in response to audience behavior and cultural context [3], [2].

VI. Conclusion

This study investigated the predictive power of intrinsic audio features for forecasting Spotify track popularity using supervised machine learning. While ensemble methods particularly the Random Forest classifier achieved moderate success in identifying popular tracks, the overall predictive accuracy remained limited, with the best models reaching approximately 81

The results underscore a fundamental limitation: audio features alone cannot fully capture the multifaceted dynamics of music popularity. Extrinsic factors such as artist reputation, marketing strategies, playlist placements, and viral social dynamics play substantial roles that remain unaccounted for in audio-only models. Nevertheless, this work demonstrates that machine learning can provide AR executives with data-driven insights to complement human intuition, potentially reducing the search space when identifying promising tracks for promotional investment.

Future research should integrate temporal, social, and metadata features alongside audio characteristics to develop more comprehensive predictive systems. By combining intrinsic sonic properties with contextual information, the music industry can move closer to reliable, scalable tools for forecasting commercial success in an increasingly data-rich streaming ecosystem.

REFERENCES

[1] S. K. Adhikary and R. Bhattacharya, "Beyond beats: A recipe to song popularity? a machine learning approach," *arXiv preprint arXiv:2403.12079*, 2024. [Online]. Available: https://arxiv.org/abs/2403.12079

[2] L. Gonzalez and P. Lee, "Decoding spotify hits: Statistical and predictive analysis of track features driving song popularity," *Academy of Marketing Studies Journal*, vol. 28, no. 1, 2024. [Online]. Available: https://www.abacademies.org/articles/decoding-spotify-hits-statistical-and-predictive-analysis-of-track-features-driving-song-popularity.pdf

[3] H. C. Chien and C. F. Chang, "Predicting music popularity using spotify and youtube features," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 9, pp. 250–259, 2022. [Online]. Available: https://www.researchgate.net/publication/363916522_Predicting_Music_Popularity_Using_Spotify_and_YouTube_Features

[4] R. R. Ghosh, M. Gupta, and S. Ray, "An analysis of classification approaches for hit song prediction using engineered metadata features with lyrics and audio features," *arXiv preprint arXiv:2301.13507*, 2023. [Online]. Available: https://arxiv.org/abs/2301.13507

[5] M. W. Dewantara, R. Pramudita, and R. Purwanto, "Predicting song popularity based on spotify's audio features: Insights from the indonesian streaming users," *Journal of Data Science and Artificial Intelligence*, vol. 4, no. 3, 2023. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/23270012.2023.2239824

[6] E. C. D'Addario, A. Piceci, and C. Santoro, "A model-based approach to spotify data analysis: a beta glmm," *PLOS ONE*, vol. 17, no. 4, 2022. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC9042099/

[7] R. Nijkamp, "Prediction of product success: Explaining song popularity by audio features from spotify data," Bachelor Thesis, University of Twente, 2018. [Online]. Available: https://essay.utwente.nl/75422/1/NIJKAMP_BA_IBA.pdf

[8] S. V. Joshi and M. Jain, "A machine learning approach using spotify data," in *Proceedings of the 13th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*. SCITEPRESS, 2024. [Online]. Available: https://www.scitepress.org/Papers/2024/133300/133300.pdf

[9] R. Bhatia, A. Kumar, and N. Ramesh, "Spotify data analysis and song popularity prediction," *SSRN Electronic Journal*, 2024. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4793176