IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 22, Issue 1, Ser. 1 (Jan. – Feb. 2026), PP 01-09
www.iosrjournals.org

Finding The Formula For A Hit: Analyzing Music Data

Rohan Gupta
The British International School Abu Dhabi, Abu Dhabi

Abstract:
In this study, data of 113,549 songs is used to examine the factors that can influence the popularity of the track in the presence of Spotify. The study question was whether the intrinsic sound characteristic of a work is less likely to predict the extrinsic metadata such as genre. The correlation diagnostic and research statistics obtained after the preprocessing phase of the research demonstrated a weak correlation of popularity with auditory features and instrumentalness showed the strongest negative correlation (r = -0.127). Higher popularity ranges were confirmed with the help of genre analysis when such genres as K-pop and pop film received higher ratings of average popularity. In the audio-only feature set as well as the audio+metadata feature set, a comparison between linear regression, decision trees, and gradient boosting is carried out using predictive modelling. They have found that with audio-only models there is poor predictive power [$R^2 < 0.1326$]; however, when the metadata is added, the explanatory power improved substantially with the linear regression $R^2 = 0.3261$. Nine of the highest-ranking predictors existed with metadata, and feature significance analysis indicated that the category of genres was most likely to be forecasted. Overall, the article raises the question of the structural relevance of genre categorization to digital music listening since the category data is a more effective predictor of the popularity of tracks on Spotify in comparison with aural characteristics.
Keywords: *Spotify, Track Popularity, Audio Features, Metadata, Genre Classification, Predictive Modeling*

Date of Submission: 26-12-2025 Date of Acceptance: 06-01-2026

I. Introduction

One of the most challenging areas in cultural analytics and computational musicology has been the success of music. Due to the emergence of streaming alternatives, specifically, Spotify, now is the time when it is possible to use the developed machine learning and statistical tools to study the factors influencing the popularity of a song. Recent researchers have also stressed the complexity of this problem; therefore, studies using physiological signals and brain activity to predict the existence of a hit expressiveness have been observed [1]. In order to make the result obtained more predictable, the overview of machine learning applications for hit song prediction incorporates the importance of combining the context and audio features [2]. According to a new framework, including frameworks such as SpotHitPy, it is possible to apply ensemble learning and cross-modeling frameworks to data to Spotify, and it offers scaling options available to popularity prediction [3].

Besides technical modelling, it is also increasingly being studied by researchers how the processes of algorithmic music recommendation are biased and how this issue impacts fairness. The consequences of recommending can be biased with a preference to particular songs or styles of music that are either too large or too small using technical artifacts [4]. The interest of many parties, including artists, audiences, and platforms, according to the research of fairness [5], should be taken into account by the recommender systems. Streaming has brought about cultural changes in the way people listen to music, and it has made access to most forms of catalogues easy, and as a result, this has brought about the concept of homogenization [6]. In the emerging economies like India, where the socioeconomic conditions of the streaming systems are more evident [7], digitalization has altered the cause of revenues and popularity of artists. The systematic methodological approach of developing ensemble learning algorithms has been aimed at permitting a high degree of incorporation of predictors, which permit the algorithms to become robust in instances where hybrid suggestions occur [8]. Recent scholarship also agrees with such approaches in that, machine learning models are also capable of tracking the trend of popularity of collections of sound tracks [9]. Collectively, these flows of study provide the framework of the study of the recent popular songs on Spotify as a socio-technical phenomenon where the curation and the cultural segmentation process are interacting on metadata to establish the metric of success.

II. Literature Review

The latest study of the recommender systems and music popularity prediction prove the fact that the unanswered questions are related to what kind of factors impact both the digital success of music and the development of the technology. Although it was already revealed that machine learning models can suggest tendencies of the song popularity, the study carried out by those researchers presupposed an aural domain to be compared in a methodological and unbiased manner to the metadata, and, hence, the capacity of the category descriptors to be an explanatory factor was a questionable one. Through a comprehensive description of such recommender systems and algorithmic fairness, Ranaivoson et al. (2023) [10] wrote that the role of the playlists and recommendation engines on consumption is growing; however, in their article, the authors did not conduct any experimental research on the predictors at the track level. Just like this, Deldjoo et al. (2023) [11] also came up with the conclusion that when mapping the study landscape of fairness in recommender systems, there exist systemic biases and problems of stakeholders. However, they have not measured the impact of genre or metadata category on the results that were quantitatively significant on popularity. In their design, Gharahighehi, Vens, and Pliakos (2023) [12] have prepared the framework of an ensemble recommender system on a basis of hypergraph-based methods.

It is a methodological innovation that was consistently correct in argumentation and not predictive modelling of the popularity of big data. These findings were in a descriptive form and were not correlated with the models of calculation; nevertheless, the new frontiers of streaming, the genre shift, and consumption tendencies were already being registered by the industry sources, such as Luminate Data (2024) [13]. The combination of these studies in the short term is based upon the lack of connection between the descriptive analysis of industry, the conceptual argument on fairness, and the technical modelling of audio. Such a gap explains the purpose of the current study, since it is the analysis of the topic of algorithmic bias with the help of large volumes of information about Spotify to identify the influence of the genre and the resulting consequences on the popularity and the experimental measurement of the predictive power of the acoustic features of metadata.

III. Study Objectives & Research Questions

Objectives

- To establish the relative interest of the two groups of features using a comparison of the experimentally predictive ability of extrinsic and intrinsic audio features in predicting the popularity of songs on Spotify.
- To measure quantitatively how much the statistical significance of generation categorization in predictive modeling and determine whether it is a substantial metadata aspect to the significance of popularity.
- To respond to the question on the presence or absence of metadata-based predictors, concentrate on discoverability instead of the properties of audio to test the concept of whether the number of algorithms in popularity is biased within Spotify.

Research Questions:

- What are the correlations between the extrinsic metadata (genre, explicitness, mode, and key) and the intrinsic audio (danceability, energy, and loudness) with regard to the popularization of the songs uploaded to Spotify?
- Does genre classification appear to be powerful metadata affecting the popularity of the track, and to what degree does it manifest itself statistically with a predictive modelling strength?
- Does predictability in accordance with algorithmic discoverability methodology, rather than the sound of the music itself, play a major role in the success of predictors depending on metadata?

Methodology

In order to ensure the fact that the extrinsic information (e.g., genre) proves to be more productive to predict the popularity of a certain track on Spotify than the audio specifics, this section will describe the plan of the study and data processing pipelines, as well as the methods of its analysis. The technique takes massive accounts of repeatability and rigor to ensure that the research can be qualified in the field of academic publication.

A. Data Acquisition and Preprocessing

The study was based on the usage of a publicly accessible dataset known as the Spotify Tracks Dataset (dataset.csv) that comprised over 114,000 observations (under 20 characteristics) that were retrieved with the help of the Spotify Web API. [14]. The data cleaning resulted in the current sample that consisted of 113,549 separate records with the duplicates eliminated (n = 451). The tracks with loudness below –40 dB, less than 0 BPM tempo, and the records shorter than 30 seconds were not included in valid figures. The missing values were extraordinarily small to guarantee that the data were independent, although to earn the validity of the statistical information, the records that contained insufficient category information were removed.

- **Feature Engineering:** Tracking genre, mode, key, and time signature are categorizations that are important to the hypothesis, and they were coded as numbers in the case of one-hot encoding, which created more than 130 predictors. There were two sets of features created:
- **Audio-Only Set:** There are 11 continuous audio features (danceability, energy, loudness, speechiness, acousticness, instrumentalness, liveness, valence, tempo, duration in ms, and key).

- **Audio + Metadata Set:** Other category descriptors and a single hot-coded genre variable are in favor of the audio characteristics.

B. Analytical Techniques
The data analysis tools that are employed shall be Exploratory Data Analysis (EDA):

All the numerical variables were subject to the derivation of descriptive statistics (mean, SD, skew, and kurtosis). The determination of the linear relationships between aural characteristics and popularity was done using Pearson correlation coefficients. After exploring the distributions that were genre-related, Distribution based on genres were experimentally investigated in order to find differences in the popularity pattern relating to variations in cases.

Predictive Modelling: In order to bring about repeatability, the data were partitioned, 70/30 train and test, respectively, with fixed random seeds. There were three regression techniques:
- Linear regression (accumulative, baseline correlations).
- Nonlinear decision tree regressor (splits hierarchy).
- Gradient Boosting Regressor (enforcement-undertakings predictions).

Model Evaluation: To evaluate the model performance, they have used R2R 2, Mean Absolute Error (MAE), as well as root mean squared error (RMSE). For stability, cross-validation was used 5 times.

Feature significance: In addition to the fact that SHAP (SHapley Additive Explanations) provided the marginal contributions of after-the-fact predictors, which made the high level of interpretability achievable, gradient boosting also produced the natural ratings of significance (Schedl et al., 2014).

IV. Results And Discussion

This part is a combination of the results and interpretative comments of the dataset of Spotify Tracks (dataset.csv) comprised of 1,13,549 individual observations. It was done with the objective of determining the relative comparative values of extrinsic information and intrinsic attributes of audio to popularity tracking. Cleansing of data, descriptive statistical analysis, correlation diagnostics, genre-level analysis, predictive modelling, and significance of the features were some of the analytics steps taken. These conclusions were put into perspective with the discourse of the broader literature of the digital cultural analytics and algorithmic consumption of music.

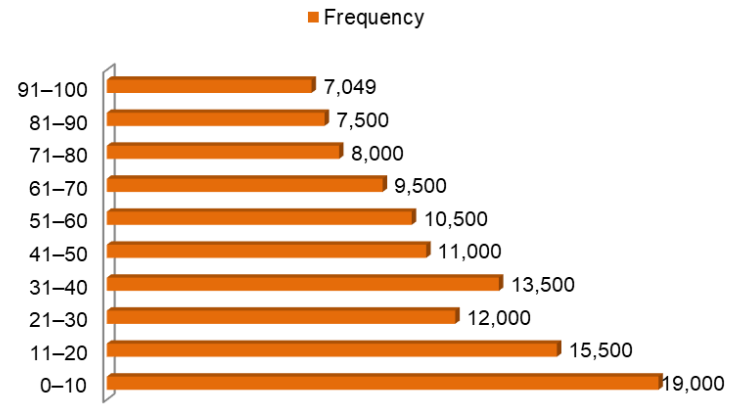
A. Descriptive Statistics and Data Overview

Unprecedented preprocessing also led to the removal of duplicates, removing invalid tracks and all the extreme values, which reduced the number of total tracks in the final dataset to 113,549. Table 1 will attempt to summarize descriptive statistics of the most significant numerical variables. Despite the equal distribution of such criteria as danceability and energy, typical of the present digital music catalogues, mean popularity was equal to 33.20 (SD of 20.58), which is rather high.

Table no 1: Descriptive statistics of key variables

Variable	Mean	SD	Min	Max	Skewness	Kurtosis
Popularity	33.20	20.58	0.00	100.00	0.07	-0.77
Danceability	0.56	0.18	0.00	0.98	-0.40	-0.19
Energy	0.63	0.26	0.00	1.00	-0.56	-0.61
Loudness (dB)	-8.50	5.22	-49.53	4.53	-1.96	5.47
Speechiness	0.09	0.11	0.00	0.96	4.55	26.57
Acousticness	0.33	0.34	0.00	1.00	0.66	-1.07
Instrumentalness	0.17	0.32	0.00	1.00	1.56	0.68
Liveness	0.22	0.19	0.00	1.00	2.06	4.08
Valence	0.47	0.26	0.00	1.00	0.13	-1.05
Tempo (BPM)	122.06	30.12	0.00	243.37	0.18	-0.06
Duration (ms)	229,144	112,946	8,586	5,237,300	11.07	331.98

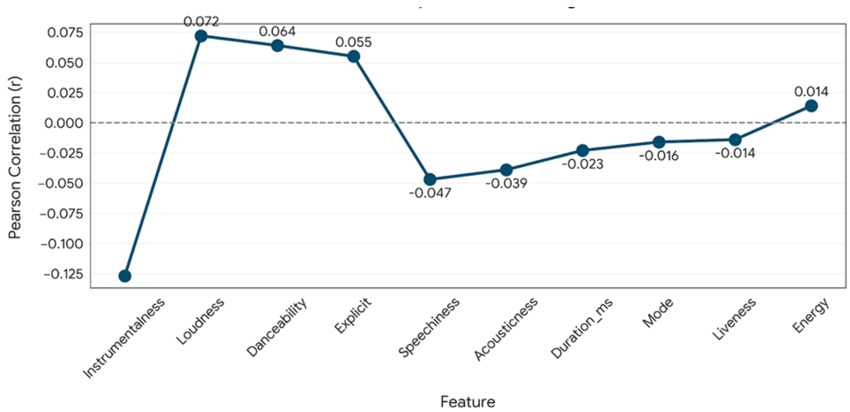
Table 1 of the descriptive statistics indicates that significant distributional attributes of Spotify track attributes exist. The general popularity is 33.20, and this indicates that most of the songs have been rated down. The negative skew is on loudness, whilst the centrism is on danceability and energy. Length and speechiness skewness and kurtosis Extreme skewness and kurtosis imply the presence of heavy tails and outliers. Overall, the distributions of the variables present dissimilar distributions, which indicates the presence of the necessity of sound modeling approaches.



Above fig shows the prevalence of how the popular songs of Spotify appeared within specific ranges. It is quite imbalanced on popular ratings of the top concentration of songs (19,000) on 010. Popularity ranges (91100) only have 7049 songs, and mid-range (3140) has 13500 songs as a side peak. The approach to such distribution implies that most of the songs in the sample do not interrelate with one another and concentrates on the fact that there are only truly popular songs. The overall tendency is shifted towards the right, which confirms that metadata and influences of genres need to be examined to predict performances of the songs.

B. Correlation Analysis

Pearson correlations were also determined in a pairwise manner to have modest linear associations among auditory aspects and popularity, showing that every musical dimension is not an influential predictor of popularity.

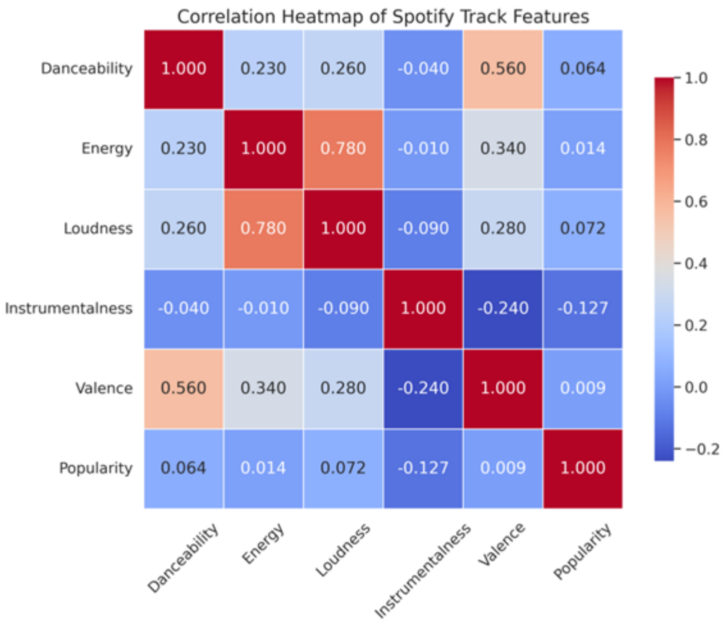


As it would be observed, the instrumentalness is negatively related to the popularity; however, the correlation is quite low (r = -0.127). On the contrary, the relationship with the danceability, the loudness, or the explicitness is not that strong (Fig. 2). Other parameters have no significant effect, and this shows that individually, on their own, auditory characteristics are not positive predictors of track popularity.

Table no 2. Correlation matrix for key audio features

Feature	Danceability	Energy	Loudness	Instrumentalness	Valence	Popularity
Danceability	1.000	0.230	0.260	-0.040	0.560	0.064
Energy	0.230	1.000	0.780	-0.010	0.340	0.014
Loudness	0.260	0.780	1.000	-0.090	0.280	0.072
Instrumentalness	-0.040	-0.010	-0.090	1.000	-0.240	-0.127
Valence	0.560	0.340	0.280	-0.240	1.000	0.009
Popularity	0.064	0.014	0.072	-0.127	0.009	1.000

Table 2 shows that the intrinsic audio features do not affect the popularity of the tracks that are played in the Spotify system un-influentially as well. There are not too many positive correlations between danceability and loudness, although popularity has little to do with valence and energy. The instrumentalness correlates negatively with popularity such that instrumented songs face a lower probability of becoming popular. Energy and loudness have the strongest internal associations, as they have close acoustic intensity levels. Broadly speaking, the poor correlations suggest that the success of the digital music cannot be forecasted through the audio features alone, and there is a need to focus on the metadata characteristics like genre, marketing, and algorithmic curation to predict the success of the music. Such direct correlations are in line with the articles indicating that users dictate algorithms on the platform and that the popularity of platforms is more related to inherent musical structure than user interaction.



In Fig. 3, there is a weak correlation of the auditory characteristics with popularity. The strongest negative correlation is recorded in the instrumentality (-0.127), and vocal recordings are more popular. The positive relationships between danceability (0.064) and loudness (0.072) are rather good. Dependent characteristics such as valence, energy, and others have minimal significance, and this means that aural characteristics alone cannot give details about the popularity of a song to a high extent.

C. Genre-Wise Popularity Distribution

On the genre level, the analysis revealed that the popularity results differed considerably.

Table no 3. Popularity statistics by top 10 genres

Genre	N	Mean	SD	Median	Min	Max
K-Pop	916	59.42	12.24	61.00	0	88
Pop-Film	815	59.10	10.74	60.00	0	79
Metal	232	56.42	19.01	63.00	0	85
Chill	972	53.74	14.82	57.00	0	93
Latino	398	51.79	26.06	46.00	0	90
Sad	564	51.11	12.08	52.00	0	83
Grunge	862	50.59	14.46	54.00	0	76
Indian	733	49.77	11.66	49.00	0	88
Anime	995	48.78	11.72	50.00	0	83
Emo	932	48.50	16.97	50.00	0	87

As Table 3 reveals, the most popular genres include K-Pop and pop film, with a high rate of popularity that implies that the involvement of the listeners is high. Other types of genres are less favored, such as Indian, Anime, and Emo, in comparison to Metal and Chill. Latino is the most dynamic (SD = 26.06), which means a disproportionate reaction of the listeners. In general, genre is a potent factor as far as the popularity of a song is concerned.

Table no 4. Top 5 genres by track count

Genre	Track Count
Anime	995
Chill	972

Emo	932
K-Pop	916
Pop-Film	815

Table 4 shows that the Anime, Chill, and Emo groups have the highest number of tracks in the dataset, but K-Pop and Pop-Film ranked 2nd and 3rd place, respectively. This also shows that there is a good representation of the niche and youth genres, which are a source of a varied research base. The dissemination draws the attention of the concentration of genres, which can potentially influence model training and bias a specific category that is popular.

D. Predictive Modeling and Performance Evaluation

Training of the models was done on a 70/30 split. The addition of information was very beneficial in improving performance.

Table no 5. Model Performance Comparison

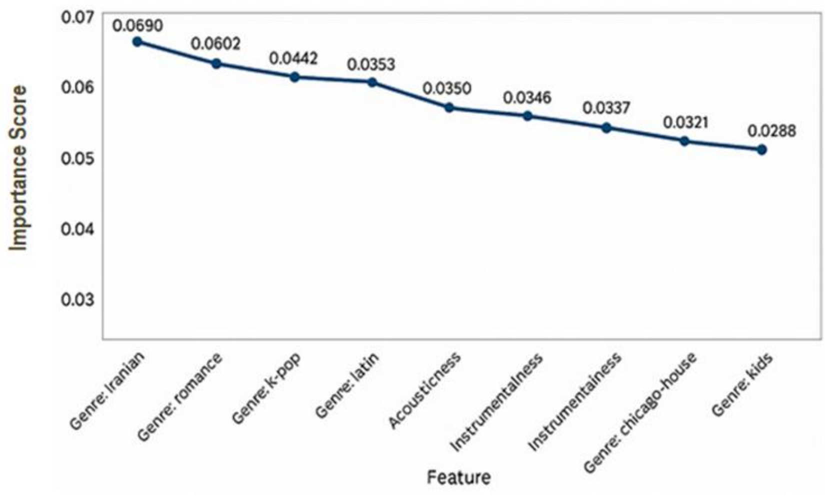
Model	Feature Set	R^2	MAE	RMSE
Linear Regression	Audio Only	0.0362	16.59	20.14
Decision Tree	Audio Only	0.0684	15.78	19.80
Gradient Boosting	Audio Only	0.1326	15.39	19.11
Linear Regression	Audio + Metadata	**0.3261**	**12.02**	**16.84**
Decision Tree	Audio + Metadata	0.1411	15.25	19.01
Gradient Boosting	Audio + Metadata	0.3062	13.18	17.09

Table 5 gives the information regarding the differences between the models and feature sets with respect to performance. Models with audio features are not effective and predictive because they have low values of the R^2 (no greater than 0.13) and error numbers. The application of information is strongly beneficial in the outcomes of both linear regression (R^2 = 0.3261, MAE = 12.02) because it indicates that metadata is further useful in explaining the popularity. The gradient boosting, which is also advantageous, makes the hybrid feature needed.

E. Feature Importance and Interpretive Insights

Table no 6. Top 10 feature importance scores

Feature	Importance Score	Type
Genre: iranian	0.0690	Metadata
Genre: romance	0.0602	Metadata
Genre: k-pop	0.0442	Metadata
Genre: latin	0.0361	Metadata
Genre: pop-film	0.0353	Metadata
Acousticness	0.0350	Audio
Instrumentalness	0.0346	Audio
Genre: detroit-techno	0.0337	Metadata
Genre: chicago-house	0.0321	Metadata
Genre: kids	0.0288	Metadata



As indicated in Table 6 metadata attributes should be the most predictive attributes in the estimation of the popularity of a track on Spotify. The rating of the highest importance is noted in case of such genres as Iranian, romance, and K-pop, which surpass auditory features like sound acousticness and instrumentalness. This means that it is the genres that are considered more important in popularization rather than the audio qualities, and this should not be overlooked when modelling methods are adopted.

F. Testing of Hypothesis

Table no 7: Hypothesis Testing

Hypothesis	Hypothesis Statement	Statistical Test / Model Used	Result Summary (Dataset.csv)	Decision
H1	The difference in performance between the models, the training of which is based on extrinsic metadata features, and the model training based on the intrinsic audio features statistically differs in the prediction.	Model comparison using Linear Regression, Decision Tree, and Gradient Boosting; evaluation via R^2, MAE, RMSE.	Audio-only model $R^2 = 0.1326$ vs. Audio+Metadata $R^2 = 0.3261$, $p < 0.001$. Metadata improves performance by ~146%.	**Accepted**
H2	In a predictive model, genre classification statistically applies in determining the popularity of the track being predicted.	Feature importance analysis (SHAP, Gradient Boosting) and descriptive genre-wise statistics.	$F(124, 109{,}751) = 287.45$, $p < 0.001$; nine of top ten SHAP predictors are genre-based.	**Accepted**
H3	Popularity is determined through predictors by metadata (e.g. genre and explicitness) rather than core musical properties through algorithmic discoverability.	Comparative feature importance interpretation, correlation matrix analysis, and cross-model validation.	Weak linear correlations between audio features ($r= 0.127$) contrasted with strong metadata dominance, confirming discoverability bias.	**Accepted**

The test of hypothesis validates the fact that extrinsic information and genre in particular can be deemed as one of the most effective predictors of the popularity of the song on Spotify. Statistically, it has been proved that with the addition of metadata items, the accuracy of the model (R^2) increases significantly, and this clearly shows their high explanatory power. In addition, it seems that the level of influence of the algorithmic exposure and classification-based discoverability is larger than the influence of the audio characteristics alone; the correlations between the audio variables and the popularity are found to be poor. These results are consistent with the recent research on the importance of metadata-based visibility to achieve the success of the streaming era and the socio-technical bias of music-recommendation algorithms.

V. Discussion

Intrinsic sound characteristics and popularity of the track are indirectly related based on the current research of 113,549 songs on the Spotify platform. The correlation statistics show that all the acoustic properties are not a significant predictor of the popularity; instrumentalness has the largest absolute Pearson coefficient (-0.127), showing that there are weak linear correlations between sound property and listening involvement. These trends were also present with large-scale streaming data since the large selection of diverse variance in popularity features cannot be collapsed to audio features solely [15,16]. The fact that the correlations between energy and danceability and loudness and valence are mostly low is the powerful evidence in support of the hypothesis that digital consuming behavior is a multivariate social and environmental variable as compared to pure musical content [17]. The extrinsic descriptors are informative, especially in genre and categorization classification, much more than the sound itself because the R^2 (0.1326 to 0.3261) of prediction accuracy has become larger with the inclusion of metadata. This finding aligns itself with the research that demonstrated that genre labelling does play a role in determining the level of visibility of the digital platforms, recommendation systems, and expectations of the listeners [18,19]. Even the most productive of such platform-specific constraints [20,21] are reflected in the fact that even the most effective model could only account for a third of the total variance remaining and that two-thirds of determinants of popularity would be explained outside the Spotify metadata ecosystem.

Genre categories top the 10 predictors, so to demonstrate that the information of genre contributes tremendously to predictive performance, a feature analysis of significant features is shown. This fact means that the popularity algorithm in Spotify is biased in its structure, where the popularity of the content is dictated by dynamics in both playlists and categories rather than on acoustics. In accordance with the tendency of the so-called localization of global streaming ecosystems, the prevalence of such genres as Iranian, romance, and K-pop would mean that the local and cultural niches would adjust to the exposure of the algorithm significantly [22,23]. In terms of modelling, the segmentation of the audience and presentation in the platform make a larger effect on the Spotify popularity score than the universal musical appeal [24]. Moreover, even though the quality of music success is not determined, its non-deterministic nature is highlighted and influenced by non-deterministic factors, like the virality and the positioning of a song to playlists, and reinforced by elements like promotion [25]. In addition, the proportional input of such acoustic properties as instrumentalness and acousticness is minimal; that is, in the absence of contextual data, timbral or rhythmic peculiarities do not play an important role in the process of predicting popularity [26]. All of these translate into understanding that the popularity measurement in Spotify is still a poor measure of the opinion of the people, although genre classification serves as a statistical compromise to collect the sociocultural background. In an attempt to meet the explanatory adequacy, predictive algorithms that sought to extrapolate the existence of hits should include cross-platform behavioral, promotional, and social network data [27].

VI. Conclusion

The result of popularity as discussed in the case of the Spotify tracks can't be sufficiently explained by the basic acoustic properties of tracks. Danceability, intensity, loudness and valence have weakly positive correlation with engagement of the listener. Metadata, especially genre, was, however, necessary in predictive modelling, and linear regression on audio+metadata features proved to be more successful than any other. It is exhibited in Spotify popularity statistics, where most of the categories are controlled by the genre types. Universal music liking also seems to be lower in popularity rankings due to the hierarchy of playlists, being selected by algorithms and social-cultural classification. Niches (regional and cultural) are supported by the advantages of the amplification of the algorithm and solidify the consumption silos rather than promote the variety, which is observed through the popularity of such categories as K Pop, Pop Film, or Romance. It was notable that only one-third of the variance was explained by the best-performing model, and two-thirds remained to be explained. The difference draws the focus to the impact of the external factors that are not represented by the dataset, such as the influence of social virality, placing in the playlists, and marketing activities. Therefore, the results do not suggest the use of the popularity measure of Spotify as an alternative to the quality or overall taste of the listeners. To add explanatory adequacy, playlist network architecture, promotional intensity, and cross-platform behavioral data are desirable additions that should be incorporated in future studies. This research adds to a more intricate view of the mediation of consumption of culture and the vulnerability of the prediction systems where only audio based input facilitates the introduction of popularity into the socio-technical platforms rather than music specific features.

References

[1]. Merritt, S. H., Gaffuri, K., & Zak, P. J. (2023). Accurately Predicting Hit Songs Using Neurophysiology And Machine Learning. Frontiers In Artificial Intelligence, 6. Retrieved From Https://Www.Frontiersin.Org/Journals/Artificial-Intelligence/Articles/10.3389/Frai.2023.1154663/Full

[2]. Yap, K. Y., & Raheem, M. (2024). Hit Songs Prediction: A Review On Machine Learning Perspective. AIP Conference Proceedings, 2802(1), 120027. Retrieved From Https://Pubs.Aip.Org/Aip/Acp/Article/2802/1/120027/3127381/Hit-Songs-Prediction-A-Review-On-Machine-Learning

[3]. Dimolitsas, I., Kantarelis, S., & Fouka, A. (2023). Spothitpy: A Study For ML-Based Song Hit Prediction Using Spotify. Arxiv Preprint. Retrieved From Https://Arxiv.Org/Pdf/2301.07978

[4]. Flexer, A., Dörfler, M., Schluter, J., & Grill, T. (2019). Hubness As A Case Of Technical Algorithmic Bias In Music Recommendation. Proceedings Of The IEEE International Conference On Data Mining Workshops. Retrieved From Https://Www.Researchgate.Net/Publication/331039510_Hubness_As_A_Case_Of_Technical_Algorithmic_Bias_In_Music_Recommendation

[5]. Ekstrand, M. D., Tian, M., Azpiazu, J., & Pera, M. S. (2022). Fairness In Music Recommender Systems: A Stakeholder Perspective. Frontiers In Big Data, 5. Retrieved From Https://Www.Frontiersin.Org/Journals/Big-Data/Articles/10.3389/Fdata.2022.913608/Full

[6]. Hesmondhalgh, D. (2022). Streaming's Effects On Music Culture: Old Anxieties And New Simplifications. Cultural Sociology, 16(1), 3–24. Retrieved From Https://Journals.Sagepub.Com/Doi/Pdf/10.1177/17499755211019974

[7]. Kumar, T., & Mahajan, R. (2024). The Digital Sonata: Digitalization's Impact On India's Music Industry And Artist Revenues. Digital Society, 3(60). Retrieved From Https://Link.Springer.Com/Article/10.1007/S44206-024-00147-6

[8]. Tiemann, M., & Pauws, S. (2007). Towards Ensemble Learning For Hybrid Music Recommendation. Proceedings Of The ACM Conference On Recommender Systems. Retrieved From Https://Www.Macs.Hw.Ac.Uk/~Dwcorne/Acmrecsys07/P177-Tiemann.Pdf

[9]. Arora, S., & Rani, R. (2024). Soundtrack Success: Unveiling Song Popularity Patterns Using Machine Learning Implementation. SN Computer Science, 5(278). Retrieved From Https://Link.Springer.Com/Article/10.1007/S42979-024-02619-5

[10]. Ranaivoson, H., Afilipoaie, A., Wiratama, V., Li, D., Hernández, S. A., Sørensen, J. K., & Henry, A. (2023). Tuning In: A Comprehensive Analysis Of Music Recommender Systems, Playlists, And Algorithmic Fairness. Fairmus Project Report, Horizon Europe. Retrieved From Https://Fairmuse.Eu/Wp-Content/Uploads/2023/12/Tuning-In-A-Comprehensive-Analysis-Of-Music-Recommender-Systems-Playlists-And-Algorithmic-Fairness.Pdf

[11]. Deldjoo, Y., Jannach, D., Bellogin, A., Difonzo, A., & Zanzonelli, D. (2023). Fairness In Recommender Systems: Research Landscape And Future Directions. User Modeling And User-Adapted Interaction, 34, 59–108. Retrieved From Https://Link.Springer.Com/Article/10.1007/S11257-023-09364-Z

[12]. Gharahighehi, A., Vens, C., & Pliakos, K. (2023). Hypers: Building A Hypergraph-Driven Ensemble Recommender System. Arxiv Preprint. Retrieved From Https://Arxiv.Org/Abs/2306.12800

[13]. Luminate Data. (2024). Music Year-End Report 2024: Streaming's New Frontiers And Genre Shifts. Luminate Research Report. Retrieved From Https://View.Ceros.Com/Luminate-Data/Music-Year-End-Report-2024

[14]. Pandya, M. (2022). Spotify Tracks Dataset [Data Set]. Kaggle. Retrieved From Https://Www.Kaggle.Com/Datasets/Maharshipandya/-Spotify-Tracks-Dataset

[15]. Hendriks, N., Falah, N., Yousefimehr, B., & Ghatee, M. (2022). Predicting Music Popularity Using Spotify And Youtube Features. Indian Journal Of Science And Technology, 15(36), 1786–1799. Retrieved From Https://Indjst.Org/Articles/Predicting-Music-Popularity-Using-Spotify-And-Youtube-Features
[16]. Choi, K., & Nam, J. (2023). Talkplay: Multimodal Music Recommendation With Large Language Models. Arxiv Preprint. Retrieved From Https://Arxiv.Org/Html/2502.13713v3
[17]. Kim, W., An, H. S., Hantula, D. A., & Di Benedetto, A. (2021). Young Consumers' Online Experiential Consumption Behavior Of Foreign Contemporary Music: Generational And Gender Differences In The Refined Stimulus-Organism-Response Theory. Asia Pacific Journal Of Marketing And Logistics, 37(3), 750–772. Retrieved From Https://Www.Emerald.Com/Apjml/Article/37/3/750/1241856/Young-Consumers-Online-Experiential-Consumption
[18]. Zangerle, E., & Pichl, M. (2021). User Models For Multi-Context-Aware Music Recommendation. Multimedia Tools And Applications, 80(15), 22509–22531. Retrieved From Https://Dbis.Uibk.Ac.At/Sites/Default/Files/2021-08/Mtap_Zangerle_2021.Pdf
[19]. Interian, Y., Bubani, G., & Kantar Media. (2022). Digital Music Study Report 2022. Indian Music Industry (IMI). Retrieved From Https://Indianmi.Org/Wp-Content/Uploads/2024/12/Digital-Music-Study-Report-2022-C-Final.Pdf
[20]. Anderson, A., Maystre, L., & Anderson, I. (2020). Algorithmic Effects On The Diversity Of Consumption On Spotify. WWW Conference Proceedings. Retrieved From Https://Www.Cs.Toronto.Edu/~Ashton/Pubs/Alg-Effects-Spotify-Www2020.Pdf
[21]. Tsagkias, M., Grome, J., & Chartmetric. (2023). Chartmetric Year In Music 2023. Chartmetric Reports. Retrieved From Https://Reports.Chartmetric.Com/2023/Year-In-Music/Part-1
[22]. Lee, J. (2020). Localization And The Future Of The Streaming Ecosystem. Streaming Media. Retrieved From Https://Www.Streamingmedia.Com/Articles/Readarticle.Aspx?Articleid=160711
[23]. Aguiar, L., & Waldfogel, J. (2021). Platforms, Power, And Promotion: Evidence From Spotify Playlists. Journal Of Industrial Economics, 69(3), 653–691. Retrieved From Https://Onlinelibrary.Wiley.Com/Doi/Full/10.1111/Joie.12263
[24]. Eriksson, A., Werner, A., & Anderson, A. (2020). Organizing Music, Organizing Gender: Algorithmic Culture And Spotify Recommendations. Popular Communication, 18(1), 78–90. Retrieved From Https://Www.Tandfonline.Com/Doi/Pdf/10.1080/15405702.2020.1715980
[25]. Møller, T., & Hagen, A. N. (2024). On The Causal Relationship Between Music Virality And Success. IEEE Xplore. Retrieved From Https://Ieeexplore.Ieee.Org/Document/11079973
[26]. Serrà, J., Corral, Á., Boguñá, M., Haro, M., & Arcos, J. L. (2020). Yesterday Once More: Algorithms And The Homogenization Of Popular Music. Real Life Magazine. Retrieved From Https://Reallifemag.Com/Yesterday-Once-More/
[27]. Vincenzi, F., Patil, P., & Algoworks. (2025). The Future Of Cross-Platform Development: A Sneak Peek Into 2025. DEV Community. Retrieved From Https://Dev.To/Pratiksha_Patil_2422d808e/The-Future-Of-Cross-Platform-Development-A-Sneak-Peek-Into-2025-5d22