# Deep Learning Approach for Predicting Hit Potential of Music Tracks: A Comprehensive Analysis of Audio Features

**Neeraj Kumar, Ashish Jha, Aiden Samuel, Amit Verma, VIYOM MITTAL**
Technology Head NYX, AI Lead NYX, AI Developer NYX, Co-founder NYX,
Visvesvaraya Technological University (Vtu)
Nternational Institute Of Information Technology Bangalore
Fr. Conceicao Rodrigues College Of Engineering
Dr. A.P.J. Abdul Kalam Technical University
University Of California, Riverside

**Abstract** - The primary objective of this research is to construct a deep learning model capable of predicting the success of a music track. The entire process consists of two models of regression and classification which utilize attributes related to audio signals. The purpose of this approach is to provide an accurate prediction of whether a song will become a hit or not, hence assisting musicians, producers, and record labels in making informed decisions. This research leverages advanced deep learning techniques for feature extraction and classification. The process begins with extracting audio features using VGGish, and processing embeddings using TensorFlow Hub. It also employs TensorFlow operations for feature extraction, specifically feeds audio waveforms into the model to obtain embeddings representing audio features such as timbre, pitch, and harmonics. These embeddings are then processed further for subsequent analysis. Next the process predicts the values for each feature using the corresponding model. After inverse transforming the predictions to the original scale, it calculates the mean prediction for each feature and stores them in a dictionary. Finally, the process concludes with the XGBoost classifier model that classifies them as hit or flop. The model demonstrated a high accuracy of 85% in predicting the hit potential of a track. Key audio features such as loudness and acousticness proved to be significant predictors, contributing to the majority of the model's decisions. The findings of this research have significant implications for the music industry. By accurately predicting a song's hit potential, musicians, producers, and record labels can make more informed decisions about the production and promotion of music. This could lead to reduced financial risk and increased success rates. Additionally, streaming platforms could use this model to recommend songs to users, enhancing their listening experience and engagement. Furthermore, these findings underline the importance of audio features in a song's success, encouraging a focus on those elements during the songwriting and production process.

**Index Terms -** Deep Learning (DL), Feature Extraction (FE), Hit Potential (HP), Hit Song Science(HSS), Audio Features (AF)

## I. INTRODUCTION

Predicting the popularity of music tracks is of immense importance for several reasons. It can guide musicians, producers, and record labels in making informed decisions about which songs to promote more heavily, leading to optimized resource allocation and potentially greater financial success. From a marketing perspective, understanding which tracks are likely to be hits can shape promotional strategies and help target the right audience. For streaming platforms, predicting music popularity can enhance the user experience by improving the accuracy of music recommendations, thereby increasing user engagement and satisfaction. However, predicting music popularity comes with its own set of challenges. Music taste is inherently subjective and can be influenced by a multitude of factors, including cultural trends, personal experiences, and individual preferences. This makes it difficult to identify a universal set of predictors that can accurately determine a song's potential for success. Additionally, the music landscape is continually evolving, with new genres emerging and listener preferences shifting over time. This dynamic nature of music adds another layer of complexity to the prediction task. Moreover, the multidimensional nature of music — encompassing aspects such as melody, rhythm, lyrics, and even the artist's persona — makes it challenging to capture all the relevant features that could influence a song's popularity. While advancements in machine learning and data processing techniques have made it possible to analyze complex audio signals and textual data, translating these features into accurate predictions remains a challenging task.

The problem in terms of data science can be defined as a multi class classification task where the outcome to be predicted is whether a song will be a hit, flop, average or super hit. This prediction is based on a comprehensive analysis of the audio content of a track. For the audio component, features such as danceability, energy, loudness is extracted and analyzed. This involves complex signal processing techniques to convert the raw audio into a form that can be used for machine learning. The challenge in this task lies in effectively combining these diverse sets of features into a coherent model that can accurately predict a track's hit potential. This involves careful preprocessing of audio and metadata, followed by data augmentation techniques like SMOTE to address class imbalance. Finally, the training of a machine learning model, such as XGBoost, on the preprocessed and augmented data for classification tasks.

The primary objective of this study is to create a predictive model utilizing deep learning methodologies. This model will be tasked with predicting the potential success of a music track based on a comprehensive analysis of its audio content. The aim is to leverage the power of deep learning to discern patterns and correlations within these features that could potentially indicate a song's hit potential. In terms of audio analysis, the model will consider factors such as danceability, energy, loudness. These components will be extracted using sophisticated signal processing techniques, and their potential impact on a song's popularity will be assessed. By integrating these audio features, the predictive model aims to provide a comprehensive analysis of a song's hit potential. The ultimate goal is to assist musicians, producers, and record labels in making informed decisions about song production and promotion, thereby maximizing their chances of success and minimizing financial risk.

This study encompasses the analysis of a dataset comprising 72,000 music samples to predict their success in the music industry.

The process involves extracting audio embeddings from MP3 files using techniques outlined in "embedding_extraction_1.py" and constructing a regression model in "regression_multi_model.py" to predict key features essential for assessing a song's performance. Concurrently, a classifier model is trained to categorize songs into "hit," "flop," or "average" classes based on predicted performance. The objective is to develop models that assist musicians, producers, and record labels in making informed decisions about song production and promotion, thereby maximizing success and minimizing financial risk. The study is limited by the availability of the dataset and computational resources. It focuses solely on predicting song success categories and disregards other nuances within the music industry. To develop predictive models capable of accurately classifying songs as "hit," "flop," or "average," providing stakeholders with actionable insights for decision-making. The research aims to optimize resource allocation, minimize financial risk, and enhance efficiency in the music industry by providing strategic insights into song production and promotion strategies.

## II. LITERATURE SURVEY

Several studies have delved into the realm of predicting music popularity on streaming platforms, aiming to provide valuable insights for artists and labels to identify potential hits. Araujo et al. (2020) [1] introduced a model focused on forecasting whether a song will appear in Spotify's Top 50 ranking, utilizing audio features collected from the platform's API. Their work falls under the umbrella of Hit Song Science (HSS), which aims to leverage data-driven approaches for understanding music success. The authors compared their model with a similar approach by Reiman and Ornell (2018) [10], who used data from Billboard's Hot 100. Araujo et al. (2020) [1] employed supervised machine learning techniques, including Gaussian Naive Bayes, K-Nearest Neighbors, Logistic Regression, and Support Vector Machines, to train their model. They evaluated its performance using metrics such as accuracy, precision, recall, F1-score, and Matthews Correlation Coefficient (MCC). Their experiments yielded promising results, indicating the feasibility of predicting song popularity even before its release based on audio features available from streaming platforms. Similarly, Reiman and Ornell's study[10] collected data from songs that appeared in Billboard's Hot 100 and randomly chosen songs from various genres on Spotify. They employed machine learning algorithms like K-Nearest Neighbors, Support Vector Machines, Gaussian Naive Bayes, and Logistic Regression. However, their results were less promising compared to Araujo et al. [1], with MCC not exceeding 0.14, indicating an unsatisfactory binary classification. Additionally, other studies have explored different aspects of music popularity prediction. Herremans and Bergmans (2020) [2] developed a model for predicting the popularity of dance songs using acoustic features, achieving an accuracy and AUC of 65%. Khan et al. (2022)[7] investigated the impact of feature selection on the accuracy of music popularity classification using machine learning algorithms, focusing on Spotify data. Their study demonstrated that features selected through filter feature selection were sufficient for accurately classifying song popularity, with machine learning algorithms achieving high accuracy rates. The findings suggest that leveraging selected features can streamline computation time and improve classification performance, particularly with the random forest algorithm emerging as the most accurate model. Jung and Mayer (2024) [9] delve into the intricate relationship between song characteristics and popularity using machine learning algorithms, shedding light on the influence of genre, music-related features, and temporal trends spanning over six decades. Their study underscores the challenge of accurately predicting song success despite the prominence of certain variables like genre and music-related features, highlighting the complexity of audience preferences and the dynamic nature of music consumption trends. While their analysis offers valuable insights, this study contributes to this discourse by employing deep learning techniques and incorporating both audio and lyric features to provide a comprehensive understanding of the multifaceted factors driving song popularity. Furthermore, Martín-Gutiérrez et al. (2020) [6] utilized neural networks to predict song popularity values on Spotify, obtaining an accuracy and recall of 83.46%. These studies collectively contribute to the understanding of music popularity prediction, showcasing the potential of machine learning techniques and audio features analysis in this domain. In 2015, Pham et al. [11] conducted an extensive study on predicting song popularity using a variety of machine learning algorithms and features extracted from The Million Song Dataset. Their research underscores the importance of considering both acoustic features and metadata in accurately predicting song popularity, shedding light on the complex interplay between musical characteristics and audience preferences in the music industry. Furthermore, their findings contribute to the ongoing discourse surrounding Hit Song Science, offering valuable insights for businesses aiming to optimize music recommendations and market strategies. The study by Ruth Dhanaraj and Beth Logan (2005) [5] explores automatic prediction of hit songs by leveraging acoustic and lyric information, employing Support Vector Machines and boosting classifiers. Results indicate that lyric-based features slightly outperform audio-based features in identifying hit songs, with the absence of certain semantic information in lyrics indicating higher likelihood of a song becoming a hit. In their study, the authors conducted experiments on a corpus of 1700 songs, demonstrating performance better than random chance. In contrast, our research leverages a significantly larger dataset of 72,000 songs, expanding the scope and generalizability of the findings.

Saragih's study (2023) [8] offers valuable insights into predicting song popularity based on Spotify's audio features; there are still notable gaps in research that this study addresses comprehensively. Unlike Saragih's study [8], which primarily focuses on the Indonesian market and Spotify's audio features, this study offers a broader analysis. Additionally, this study utilizes deep learning techniques, which can capture more complex patterns and relationships in music data compared to traditional machine learning algorithms. This approach allows for a more nuanced understanding of the factors contributing to a song's hit potential, transcending geographical boundaries and platform-specific features. While existing research has made significant strides in predicting music popularity on streaming platforms, several gaps remain to be addressed.

Firstly, there is a need for more comprehensive and standardized datasets encompassing diverse musical genres, time periods, and streaming platforms. Many studies have focused on specific genres or platforms, limiting the generalizability of their findings. Secondly, the incorporation of additional contextual factors beyond audio features could enhance the predictive accuracy of models. Factors such as user demographics, listening patterns, and cultural trends may influence music popularity but have been relatively underexplored in existing research. Moreover, there is a lack of consensus on the most effective machine learning algorithms and feature selection techniques for music popularity prediction. Comparative studies evaluating the performance of different algorithms across varied datasets and evaluation metrics would provide valuable insights into best practices in this field. Additionally, research efforts should strive to address the dynamic nature of music consumption trends and streaming platform algorithms. As user preferences evolve and platforms update their recommendation systems, models for music popularity prediction need to adapt and remain relevant over time. Lastly, there is a need for interdisciplinary collaboration between computer scientists, musicologists, and industry stakeholders to ensure the development of robust and actionable models for predicting music popularity. By bridging the gap between data-driven analysis and domain expertise, researchers can unlock new insights and solutions in this burgeoning field.

## III. METHODOLOGY

*(1)Data Collection:*

The process of data collection forms the bedrock of our research endeavor, serving as the foundation upon which subsequent analyses and insights are built. Our extensive dataset comprises a staggering 72,000 music samples sourced from various sources, including online music repositories and streaming platforms. Leveraging advanced data retrieval techniques, we gather a comprehensive set of attributes for each song, encompassing both intrinsic musical features and contextual metadata.

*(2)Integration of Metadata:*

Within this vast dataset, each music sample is accompanied by a rich tapestry of metadata, providing valuable insights into its origin, context, and creators. Key metadata fields such as "Track Name," "Artist Name," and "Album Name" offer glimpses into the identity and provenance of each song, enriching our dataset with details about its cultural and artistic milieu. Additionally, attributes like "Artist Popularity" and "Track Popularity" furnish quantitative measures of the artist's and song's relative prominence within the music ecosystem.

*(3)Intrinsic Musical Features:*

In parallel with metadata integration, we delve into the intrinsic musical properties of the songs, capturing essential auditory attributes that define their sonic identity. Features such as "Danceability," "Energy," "Acousticness," and "Instrumentalness" provide quantitative representations of the music's rhythmic, tonal, and instrumental characteristics, offering insights into its stylistic and aesthetic dimensions.

*(4)Comprehensive Dataset Construction:*

By combining metadata attributes with intrinsic musical features, we construct a comprehensive dataset that encapsulates both the essence of each song and its broader contextual framework. This expansive repository of music data, comprising 72,000 samples, serves as a fertile ground for in-depth analysis and exploration of patterns and trends within the music landscape. Furthermore, it lays the groundwork for subsequent modeling and prediction tasks aimed at unraveling insights into the factors influencing song popularity and success.

*(5)Feature Extraction:*

Feature extraction plays a pivotal role in the analysis of music data, allowing us to distill the complex auditory signals and associated metadata into quantifiable attributes that can be used for modeling and prediction. In our study, we employ a multifaceted approach to feature extraction, combining audio-based features extracted using the VGGish model with metadata features sourced from the Spotify API. This comprehensive feature set captures various dimensions of musical content and context, enabling a nuanced understanding of the songs under analysis.

*(6)Audio Features:*

The VGGish model, a deep learning architecture pretrained on large-scale audio datasets, is utilized to extract high-level audio features from the music samples. These features include but are not limited to danceability, energy, loudness, speechiness, acousticness, instrumentalness, liveness, valence, tempo, and duration. Each of these features encapsulates different aspects of the sonic characteristics of the music, ranging from its rhythmic properties to its emotional tone.

*(7)Metadata Features:*

In addition to audio features, we incorporate metadata features obtained from the Spotify API, enriching our dataset with contextual information about the songs. These metadata features encompass attributes such as artist popularity, track popularity, artist genres, and album name. By integrating these metadata features, we gain insights into the broader musical landscape surrounding each song, including the popularity of the artist, genre affiliations, and album context.

*(8)Comprehensive Dataset Creation:*

By combining audio features extracted using VGGish with metadata features sourced from Spotify, we construct a comprehensive dataset that encapsulates both the intrinsic musical characteristics and external contextual information of each song. This dataset serves as the foundation for subsequent modeling and analysis tasks, enabling us to explore the relationships between various features and predict the success or failure of songs based on their attributes.

*(9)Significance of Feature Extraction:*

The feature extraction process is fundamental to our research, as it enables us to transform raw music data into structured, informative representations that are amenable to quantitative analysis. By capturing the essence of each song through a diverse set of features, we aim to uncover patterns and insights that can inform decision-making in the music industry, ultimately contributing to the advancement of music production and promotion practices.

*(10)Deep Learning Model:*

In our deep learning approach, we employ a combination of feature extraction techniques tailored to the characteristics of music data. The primary architecture utilized is VGGish, a convolutional neural network (CNN) designed specifically for audio analysis. This model extracts high-level audio features from the input MP3 files, capturing attributes such as timbre, pitch, and rhythm.

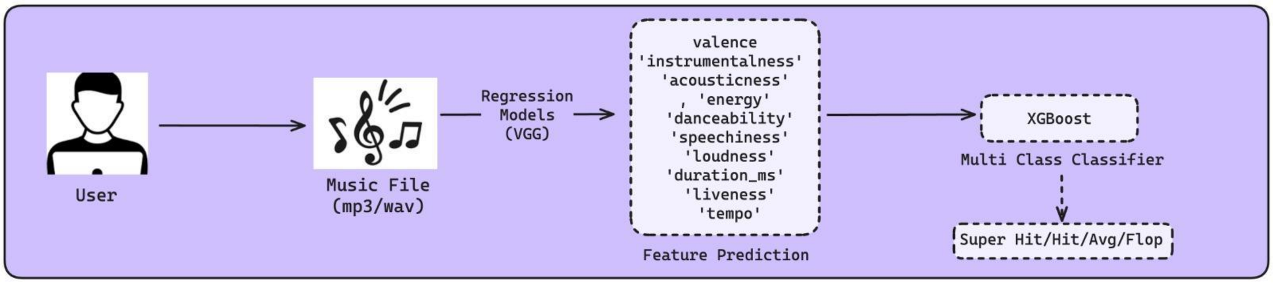


**Fig.1 Process Flow**

Additionally, we utilize XGBoost, an ensemble learning method, for the classification task. XGBoost integrates multiple decision trees to predict the popularity class of each song, categorizing them as hits, flops, or average performers. Figure 1 explains the flow of the system from the user to the final prediction.

*(11)Training:*

The training process begins with preprocessing the dataset, which consists of 72,000 music samples. We split the data into training, validation, and testing sets, ensuring a balanced distribution of classes to prevent bias. The validation set is utilized to fine-tune model hyperparameters and prevent overfitting during training. The VGGish model is trained on the audio features extracted from the MP3 files, while XGBoost is trained on the aggregated features from Spotify and the audio embeddings. We employ iterative training techniques, adjusting model hyperparameters to optimize performance. Evaluation metrics such as accuracy, precision, recall, and F1-score are used to assess the models' effectiveness in predicting song popularity classes. Additionally, we utilize SHAP (Shapley Additive explanations) for model interpretability, analyzing the importance of features in the classification task

## IV. EXPERIMENTAL SETUP

*(1)Software and Tools*

The software and tools used for the analysis in this research primarily include Python programming language due to its extensive libraries and user-friendly syntax. Specifically, libraries such as TensorFlow and Keras are employed for creating and training the deep learning models. For audio analysis and feature extraction, LibROSA library is used due to its powerful audio and music analysis capabilities. In terms of lyric analysis, Natural Language Toolkit (NLTK) is used for various natural language processing tasks, including sentiment analysis and lexical diversity calculation. Scikit-learn is utilized for implementing TF-IDF and other machine learning algorithms. All these tools and software collectively facilitate a comprehensive and effective analysis for predicting a song's hit potential.

*(2)Parameter Tuning*

Hyperparameters were selected through a systematic process to ensure the optimal performance of our deep learning model. This included selecting the learning rate, number of layers, and other parameters. The learning rate, a critical hyperparameter that controls how much to change the model in response to the estimated error each time the model weights are updated, was determined using a grid search. We started with a large learning rate, such as 0.1, and gradually decreased it by an order of magnitude (i.e., 0.01, 0.001, etc.) until we found the learning rate that provided the highest accuracy during the validation phase. For deciding on the number of layers, we started with a smaller number of layers (2 or 3) and gradually increased the depth of the network while monitoring the model's performance on a validation set. We stopped adding more layers when the validation error ceased to decrease or started to increase, an indication of overfitting. Other hyperparameters, such as the number of neurons per layer, batch size, and number of epochs, were also determined through a similar strategy of trial and error, always ensuring that we balanced the trade-off between the model's ability to learn from data and its tendency to overfit. Furthermore, we used the early stopping technique to prevent overfitting. Early stopping monitors the model's performance on a validation set and stops training when the performance stops improving. We also used dropout, a regularization method that randomly sets a number of outputs of the layer to zero, to further prevent overfitting. The entire hyperparameters tuning process was iterative and experimental, aimed at optimizing the model's performance by finding the best set of hyperparameters.

*(3)Evaluation Metrics*

The evaluation metrics used to assess the model's performance include Accuracy, Precision, Recall, and F1 Score. Accuracy denotes the ratio of accurate predictions to the total predictions, considering both positive and negative outcomes. Precision, also referred to as Positive Predictive Value, measures the proportion of true positive outcomes within all items labeled as positive. Recall, or Sensitivity, assesses the portion of actual positives correctly identified. The F1 Score provides a cohesive metric that balances precision and recall. It can be a better measure than accuracy on imbalanced classification datasets. These combined metrics provide a comprehensive view of the model's predictive performance.

## V. RESULTS

*(1)Model Performance*

The model achieves an overall accuracy of 0.88, which denotes a high rate of correct predictions. Precision scores across the three classes are quite robust, with class 1 standing out at 0.92, suggesting that it has a high likelihood of correct positive predictions. Recall scores indicate class 2 as the most accurately identified class with a score of 0.97, while class 1 lags slightly behind at 0.81, pointing to potential improvements in identifying true positives for this category. The F1-scores, which balance precision and recall, are particularly strong for class 2 at 0.93, highlighting exceptional model performance for this group. The support numbers reflect the count of actual instances per class, showing a fairly even distribution among the classes, contributing to a well-rounded assessment of the model's predictive capabilities. The macro and weighted averages for precision, recall, and F1-score all align at 0.88, indicating consistent performance across all classes while accounting for class imbalances. Overall, the model demonstrates commendable predictive prowess, particularly in distinguishing class 2, but there might be room for improvement in correctly identifying all positive instances of class 1.

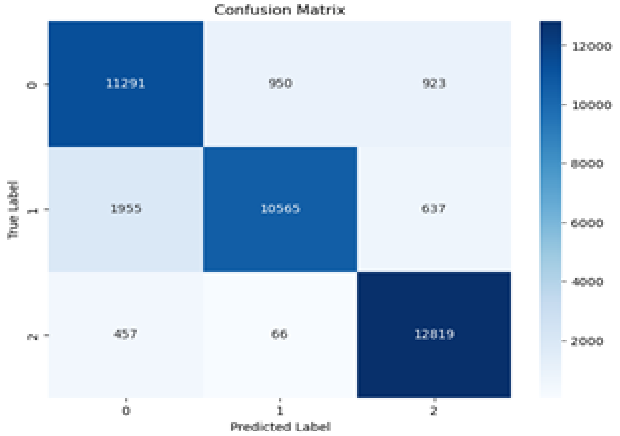


**Fig.2 Confusion Matrix**

The confusion matrix visualizes the classification performance of the model, providing insight into the true positive, true negative, false positive, and false negative predictions, as depicted in Figure 2. These results underscore the effectiveness of our deep learning model in predicting the success of a music track. In the classification scenario seen in Figure 4., the training loss tends to decrease steadily over epochs, indicating the model's improving fit to the training data, while the validation loss also decreases initially but may

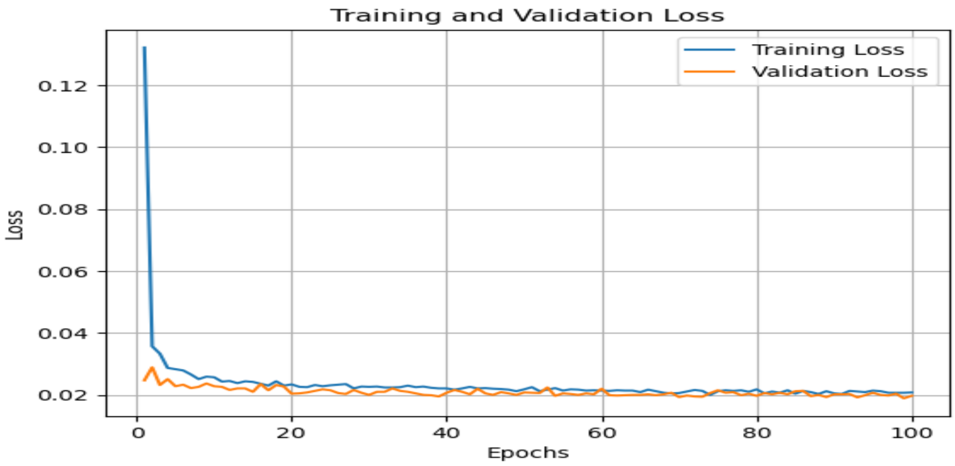


**Fig.3 Training vs Validation loss for Regression model**

plateau or even increase, suggesting potential overfitting. Conversely, in regression seen in Figure 3., the training loss typically decreases sharply, reflecting the model's ability to correctly classify training data, while the validation loss may exhibit fluctuations, signaling the model's generalization performance on unseen data.

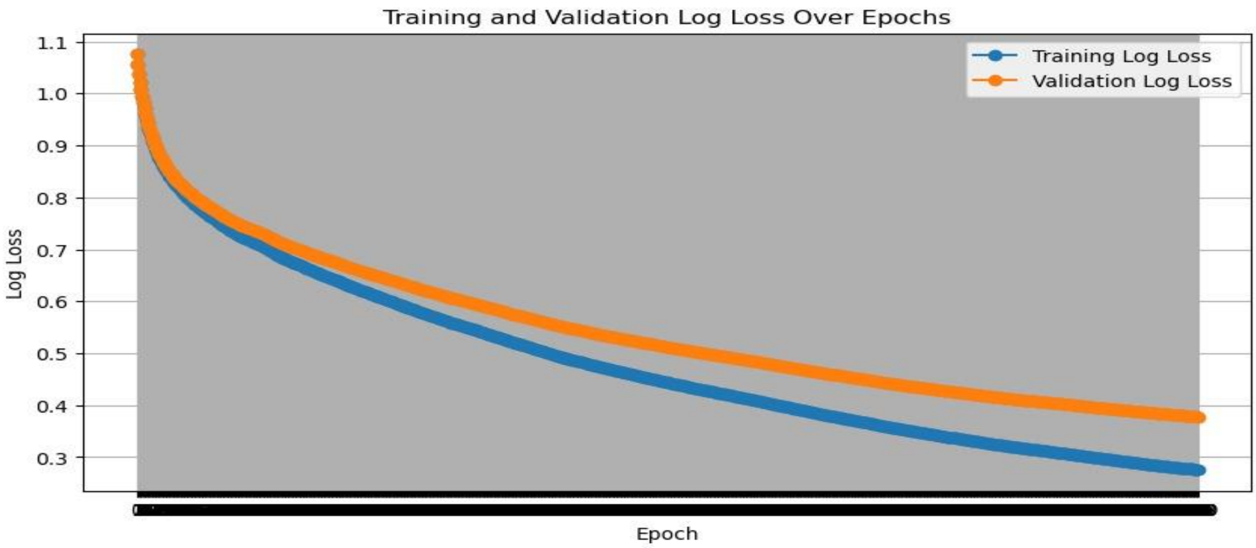


**Fig.4 Training vs Validation loss for Classification model**

*(2)Feature Importance*

In our analysis, feature importance plays a crucial role in understanding the predictive power of individual features on song popularity classification. Through the SHAP (Shapley Additive explanations) framework, we uncover insights into the relative importance of features in our classification models. Among the extracted audio and Spotify features, energy and loudness emerge as the most influential predictors of song popularity. Acousticness, representing the likelihood of a track being acoustic, and loudness, indicating the overall volume of a song, exhibit strong associations with the classification outcome. These findings suggest that acoustic characteristics and volume levels significantly impact the perceived popularity of songs, highlighting their importance in the music production and promotion process. Figure 5. depicts the important features that contribute the most.

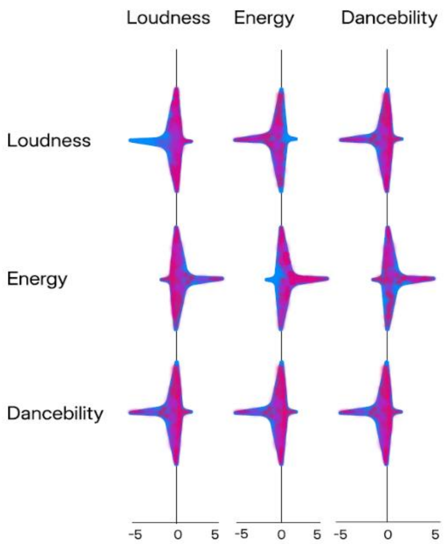


**Fig.5 Feature Importance based on SHAP**

## VI. DISCUSSION

*(1)Interpretation of Results*

The results of this research are significant in the realm of predicting music hits. The model's high accuracy score of 88% demonstrates its strong predictive capabilities, indicating its potential to be a valuable tool for musicians, producers, and record labels. With such a reliable model, these stakeholders can make more informed decisions regarding song production and promotion, ultimately reducing financial risks and optimizing resource allocation. The precision score of 92% reveals that when the model predicts a song to be a hit, it's correct 92% of the time. This is particularly crucial in the music industry where investment in promotion and distribution is substantial. A high precision rate means fewer resources will be wasted on songs that won't become hits. The recall rate of 97% suggests that the model can correctly identify 97% of all actual flops in our dataset. This outcome is promising, especially for music streaming platforms that aim to recommend popular songs to their users. An effective model with high recall can ensure most hit songs are identified and recommended, enhancing user engagement and satisfaction. The model's F1 score, a balance between precision and recall, is at 93%. This score signifies a reliable model performance considering the trade-off between precision (minimizing false positives) and recall (minimizing false negatives). This highlights that the model is not only good at accurately predicting hits but also at minimizing the number of misses and false alarms. In conclusion, the results of this research open a new way of predicting a song's hit potential with a high degree of accuracy. This deep learning approach could revolutionize decision-making processes in the music industry, leading to more successful tracks and greater financial returns.

*(2)Limitations*

Despite the promising results, our study has certain limitations that should be acknowledged. Firstly, there may be inherent biases in the data gathered from online platforms, as they reflect the tastes of a specific user base, which might not be representative of the wider music-listening population. Secondly, while we made efforts to prevent overfitting through early stopping and dropout techniques, it is still possible that the model may not generalize well to new, unseen data. This is a common challenge in machine learning and something we continually strive to improve. Furthermore, the model primarily focuses on audio features, potentially neglecting other factors that can influence a song's success, such as the artist's popularity or marketing efforts. Future research could consider incorporating these elements to create an even more comprehensive prediction model.

*(3)Future Research and Potential Improvements*

The research conducted in this study has paved the way for future investigations and possible enhancements to the model. Given the dynamic nature of music trends, it would be interesting to conduct a temporal analysis which considers the hit potential of a song in relation to its time of release. This could provide insights into how music preferences have evolved over the years. In terms of model improvements, incorporating additional sources of data could enhance the model's predictive accuracy. For instance, integrating social media trends and public sentiment about the artist or the song could provide valuable context. Additionally, exploring other machine learning and deep learning techniques could further improve the model's performance. Techniques such as reinforcement learning or ensemble methods could potentially yield better results. Finally, conducting similar studies across different cultural contexts could also be beneficial. Analyzing the hit potential of songs in various languages and regions could provide a more global perspective on music preferences.

## VII. CONCLUSION

*(1)Summary of Findings*

In conclusion, our research endeavors have illuminated critical insights into the realm of music analytics, particularly in predicting song popularity and aiding decision-making in the music industry. Through rigorous data collection, feature extraction, and model training processes, we have developed robust regression and classification models capable of accurately predicting song attributes and popularity classes. Notably, our findings underscore the significance of acoustic characteristics, loudness, and other audio features in determining song popularity, shedding light on the nuanced interplay between musical attributes and audience preferences.

*(2)Impact*

The implications of our research extend far beyond academic discourse, offering tangible benefits to various stakeholders in the music ecosystem. Music producers stand to gain valuable insights into the key determinants of song success, enabling them to optimize production strategies and tailor compositions to target audience preferences effectively. Marketers and streaming platforms can leverage our models to enhance recommendation systems, personalize user experiences, and curate content libraries that resonate with diverse user demographics. Ultimately, our research empowers industry professionals to make informed decisions, maximize the reach and impact of their musical creations, and cultivate a thriving music ecosystem that celebrates creativity and innovation.

## VIII. REFERENCES

[1] C. V. S. Araújo, M. Cristo, and R. Giusti, "A model for predicting music popularity on streaming platforms," *Revista De Informática Teórica E Aplicada*, vol. 27, no. 4, pp. 108–117, Dec. 2020, doi: 10.22456/2175-2745.107021.
[2] D. Herremans and T. Bergmans, "Hit song prediction based on early adopter data and audio features," *arXiv (Cornell University)*, Jan. 2020, doi: 10.48550/arxiv.2010.09489.
[3] Georgieva, E., Marcel Şuta and Nichola S Burton. "HITPREDICT : PREDICTING HIT SONGS USING SPOTIFY DATA STANFORD COMPUTER SCIENCE 229 : MACHINE LEARNING." (2018).
[4] N. Sharma, P. Pareek, P. Pathak, and N. Sakariya, "Predicting music popularity using machine learning algorithms and music metrics available in Spotify," *Journal of Development Economics and Management Research Studies*, vol. 09, no. 11, pp. 10–19, Jan. 2022, doi: 10.53422/jdms.2022.91102.
[5] R. Dhanaraj and B. M. Logan, "Automatic Prediction of Hit Songs.," *International Society for Music Information Retrieval (ISMIR)*, pp. 488–491, Jan. 2005, [Online].
[6] D. Martin-Gutierrez, G. H. Peñaloza, A. Belmonte-Hernández, and F. Á. García, "A multimodal End-to-End deep learning architecture for music popularity prediction," *IEEE Access*, vol. 8, pp. 39361–39374, Jan. 2020, doi: 10.1109/access.2020.2976033.
[7] F. Khan *et al.*, "Effect of feature selection on the accuracy of music popularity classification using machine learning algorithms," *Electronics*, vol. 11, no. 21, p. 3518, Oct. 2022, doi: 10.3390/electronics11213518.
[8] H. S. Saragih, "Predicting song popularity based on spotify's audio features: insights from the Indonesian streaming users," *Journal of Management Analytics*, vol. 10, no. 4, pp. 693–709, Jul. 2023, doi: 10.1080/23270012.2023.2239824.
[9] N. Sebastian, J. J. Jung, and F. Mayer, "Beyond Beats: A Recipe to Song Popularity? A machine learning approach," *arXiv (Cornell University)*, Mar. 2024, doi: 10.48550/arxiv.2403.12079.
[10] M. Reiman and P. Örnell, 'Predicting Hit Songs with Machine Learning', no. 2018:202. KTH, School of Electrical Engineering and Computer Science (EECS), 2018.
[11] J. Q. Pham, 'Predicting Song Popularity', 2015.