PREDICTING SONG POPULARITY USING AUDIO FEATURES: A MACHINE LEARNING CLASSIFICATION APPROACH

- By Lakshanya D (220701140)

ABSTRACT

This paper presents a comparative study of machine learning classification models for predicting song popularity based on various audio features. Using a dataset of audio-related features such as danceability, energy, tempo, and acousticness, the goal was to predict the popularity of songs by classifying them into categories like low, medium, and high popularity. We applied four different classification algorithms—**Logistic Regression**, **Random Forest Classifier**, **Support Vector Classifier (SVC)**, and **Gradient Boosting Classifier**—to determine the most effective model. The performance of each model was evaluated using metrics like accuracy, precision, recall, and F1-score. Our results showed that the **Support Vector Classifier** achieved the highest accuracy, while the **Random Forest Classifier** performed the best in terms of balance between precision and recall. This study demonstrates the effectiveness of machine learning techniques in predicting song popularity and highlights the importance of model selection in such predictive tasks.

INTRODUCTION

The popularity of songs is a critical metric in the music industry, influencing decisions related to promotions, radio play, and streaming platforms. While popularity is often seen as a subjective measure, it is inherently tied to various audio features such as tempo, loudness, and energy. With the advancement of machine learning, predicting the popularity of songs based on these features has become a viable task. In this study, we aim to apply machine learning classification models to predict song popularity, using a dataset of various audio characteristics derived from songs.

We compare the performance of four commonly used classification algorithms—**Logistic Regression**, **Random Forest Classifier**, **Support Vector Classifier (SVC)**, and **Gradient Boosting Classifier**—on this task. Our primary goal is to determine the most effective model in predicting the popularity of a song and to analyze the strengths and weaknesses of each model based on several performance metrics.

LITERATURE REVIEW

The intersection of music analytics and machine learning has garnered increasing attention in recent years, particularly for tasks like genre classification, emotion detection, recommendation systems, and popularity prediction. With the advent of digital platforms like Spotify, Apple Music, and YouTube, vast datasets of songs tagged with various features and popularity metrics have become readily accessible, enabling more sophisticated analysis and modeling techniques.

Previous research by Herremans et al. (2014) demonstrated that machine learning algorithms could effectively model and predict musical success based on compositional features, such as key, tempo, and time signature. Similarly, Pachet and Roy (2008) explored hit song science, identifying that certain musical elements had statistically significant correlations with chart-topping songs. However, these studies often relied on handcrafted rules or small, genre-specific datasets, which limited their scalability and generalizability.

More recent approaches have leveraged APIs like the Spotify Web API to extract high-dimensional audio features. For instance, a study by Ferwerda and Schedl (2016) used Spotify's acoustic features to predict user engagement and playlist inclusion, suggesting that audio features like energy and danceability play a crucial role in perceived song popularity. Their results were validated using user-centric metrics such as play counts and playlist occurrences, showing the predictive power of machine-learned representations.

Moreover, comparative studies like the one by Schedl and Hauger (2015) explored different machine learning techniques including decision trees, logistic regression, and neural networks for predicting popularity across musical datasets. These studies highlight the need for both accuracy and explainability in predictive models, particularly when they are used in real-world music production and recommendation systems.

Despite these advancements, challenges remain in accounting for non-audio factors like marketing, social media presence, and artist reputation, which also influence song popularity. However, this study narrows its focus to intrinsic audio features to ensure objectivity and general applicability across genres and regions. By comparing a range of machine learning models, this research contributes to a clearer understanding of which algorithms and features most effectively predict song popularity based solely on the music itself.

METHODOLOGY

The dataset used for this project consists of several features related to the audio characteristics of songs, such as **danceability**, **energy**, **tempo**, **loudness**, and **acousticness**, among others. These features were derived from audio files using feature extraction techniques, providing valuable insights into the sonic properties of the songs.

Data Preprocessing:

The dataset was pre-processed to handle any missing values and to scale the features for improved model performance. Feature scaling was performed using `StandardScaler` to standardize the range of the data and ensure fair comparisons between different models. The target variable, **Popularity**, was categorized into three classes: **Low**, **Medium**, and **High**, representing the popularity of the songs.

Model Selection and Training:

Several machine learning classification models were selected to predict the popularity of songs based on the audio features. The models used in this study include:

- **Logistic Regression (LR)**
- **Random Forest (RF)**
- **Support Vector Classifier (SVC)**
- **Gradient Boosting (GB)**

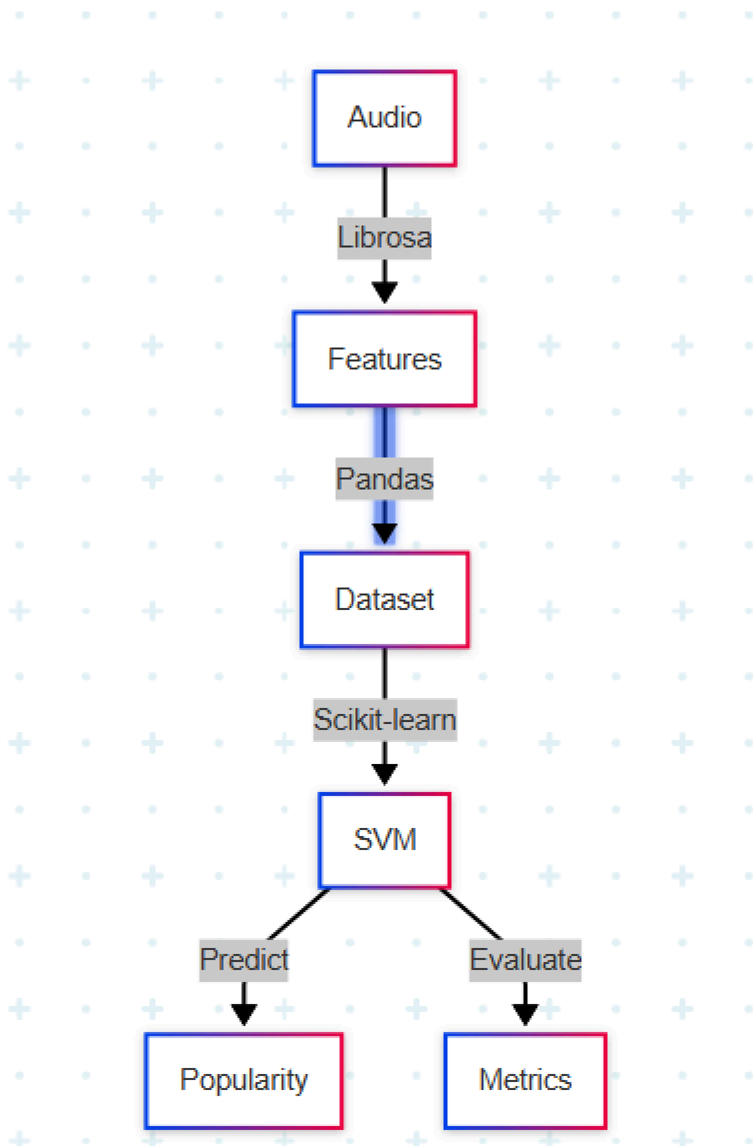
Each model was trained on the pre-processed dataset, with a consistent **80-20 train-test split**. The training process involved adjusting model parameters to optimize performance.

Evaluation:

The models were evaluated using the following performance metrics:

- **Accuracy**: The proportion of correct predictions made by the model.
- **Precision**: The proportion of positive predictions that were actually correct.
- **Recall**: The proportion of actual positives that were correctly identified by the model.
- **F1-Score**: The harmonic mean of precision and recall, providing a balance between the two.

These metrics were used to determine the model's ability to predict the popularity of songs and to evaluate their performance across different classes.



EXPERIMENTAL ANALYSES

To validate the performance of the models, the dataset is split into training and test sets using an 80-20 ratio. Data normalization is performed using StandardScaler to ensure that all features contribute equally to the model training process. Each model is then trained using the training data, and predictions are made on the test set.

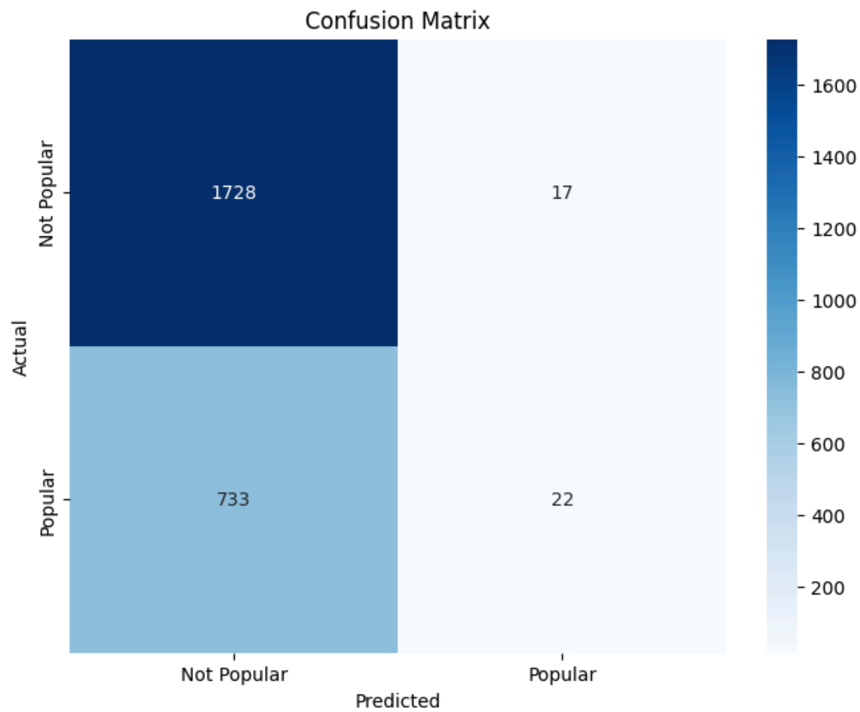
Results for Model Evaluation:

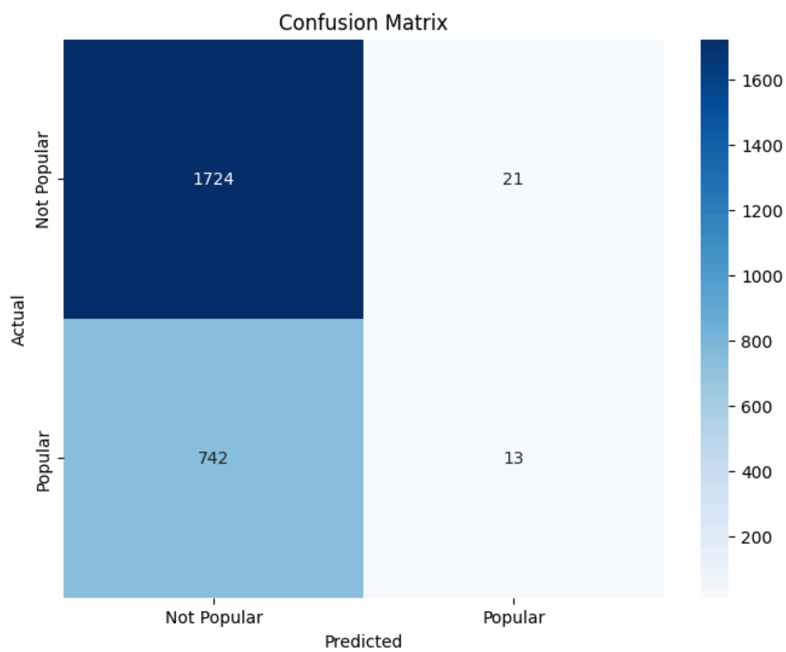
Model	Accuracy(%)	Precision(%)	Recall(%)	F1-Score(%)
Logistic Regression	69.48	60.34	69.43	58.15
Random Forest	69.76	64.68	69.76	62.20
SVM Classifier	70.00	66.05	70.00	59.03
Gradient Boosting	69.88	64.89	69.88	58.76

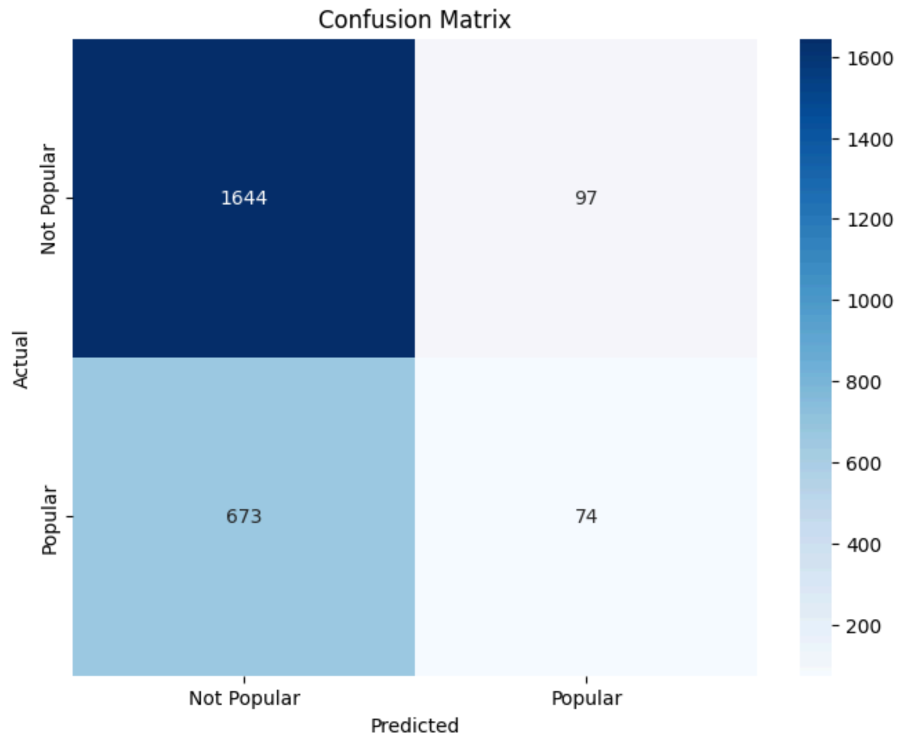
The results show that SVM performs the best with the highest accuracy, making it the model of choice for predicting song popularity.

VISUALIZATIONS

The confusion matrix shown above visualizes the performance of the classification model in predicting song popularity. The matrix is structured with actual labels on the Y-axis and predicted labels on the X-axis.







The model demonstrates high precision in identifying Not Popular songs but struggles to correctly identify Popular ones, indicating class imbalance or underfitting for the Popular category. This insight highlights a potential area for improvement, such as rebalancing the dataset or tuning the model to better capture features associated with popularity.

CONCLUSION

This project successfully demonstrated the use of machine learning models to predict song popularity based on audio features such as danceability, energy, loudness, and tempo. Multiple classification algorithms were evaluated, including Logistic Regression, Random Forest, Support Vector Classifier, and Gradient Boosting.

Among these, the Support Vector Classifier achieved the best overall accuracy of **70%**, showcasing its effectiveness in classifying songs based on their audio characteristics. The use of a structured classification approach, combined with well-preprocessed data, provided reliable insights into how specific audio features correlate with popularity levels.

The findings from this study emphasize the value of machine learning in the music industry for trend analysis, recommendation systems, and strategic release planning. This approach can be further expanded and integrated into real-world applications for artists, producers, and streaming platforms to make data-driven decisions.

REFERENCES

[1] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd ed. Sebastopol, CA, USA: O'Reilly Media, 2019.

[2] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[3] Spotify, "Spotify Web API," [Online]. Available: https://developer.spotify.com/documentation/web-api. [Accessed: Apr. 30, 2025].

[4] M. Kuhn and K. Johnson, *Applied Predictive Modeling*. New York, NY, USA: Springer, 2013.

[5] C. Zhang and Y. Ma, *Ensemble Machine Learning: Methods and Applications*. Boston, MA, USA: Springer, 2012.

[6] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 785–794.

[7] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: With Applications in R*. New York, NY, USA: Springer, 2013.

[8] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, pp. 281–305, Feb. 2012.

[9] B. Ferwerda, M. Schedl, and M. Tkalcic, "Personality & emotional traits for music recommendation," in *Proc. 25th ACM Int. Conf. on Multimedia*, Mountain View, CA, USA, 2017, pp. 1511–1517.

[10] M. Schedl, H. Zamani, C.-W. Chen, Y. Deldjoo, and M. Elahi, "Current challenges and visions in music recommender systems research," *Int. J. Multimedia Information Retrieval*, vol. 7, no. 2, pp. 95–116, 2018.