IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 13, Issue 11, November 2025

DOI: 10.17148/IJIREEICE.2025.131117

Song Skip Prediction Using XGBoost: A Machine Learning Approach for Music Recommendation Systems

Surya Sivakumar[1], Koduri Pranav[2], Tharun kumaar SD[3], Dr G.Paavai Anand[4]

Student, Department of CSE, SRM Institute of Science and Technology, Chennai, India[1]

Student, Department of CSE, SRM Institute of Science and Technology, Chennai, India[2]

Student, Department of CSE, SRM Institute of Science and Technology, Chennai, India[3]

Assistant Professor (Sr.G), Department of CSE, SRM Institute of Science and Technology, Chennai, India[4]

Abstract: With the rapid evolution of digital music platforms like Spotify and Apple Music, understanding user behavior has become crucial for enhancing music recommendation systems. One key behavioral indicator is song skipping, which reflects user preferences and engagement levels. This study introduces a Song Skip Prediction System using the XGBoost (Extreme Gradient Boosting) algorithm to predict whether a user will skip a song based on playback and contextual features. The dataset, comprising 149,860 interaction records, was preprocessed through encoding, scaling, and feature selection to ensure data consistency and accuracy.

The model achieved an impressive accuracy of 97.27% and an ROC-AUC score of 0.9770, outperforming traditional ensemble methods. Evaluation metrics such as the confusion matrix and ROC curve confirmed its strong discriminative performance. To prevent overfitting and data leakage, techniques like cross-validation and regularization were employed. The trained model was deployed using a Flask backend with a React-based frontend, allowing real-time skip predictions through a user-friendly interface. Overall, this work demonstrates how XGBoost can effectively model user listening behavior, offering a scalable foundation for intelligent and personalized music recommendation systems.

Keywords: Music recommendation, user behavior, song skip prediction, XGBoost, machine learning, user engagement, Flask, React, real-time prediction, personalization

I. INTRODUCTION

In recent years, music recommendation platforms have revolutionized how users interact with digital music libraries. These platforms generate enormous amounts of playback data every second including song duration, listening time, device type, and skip behavior providing valuable opportunities for data-driven personalization. Among these interactions, the phenomenon of song skipping serves as a critical behavioral indicator of user engagement and satisfaction. Predicting whether a user will skip a song can enhance recommendation systems, improve playlist sequencing, and contribute to a more personalized listening experience.

This research is created to predict when users are about to skip a song, based on playback data similar to real world Spotify interactions. The study employs the XGBoost (Extreme Gradient Boosting) algorithm, chosen for its reliability and strong performance in handling classification problems. Using over 149,000 playback records, the model explores the connection between audio features and listening contexts to estimate the probability of a skip event. The insights derived from this system can help streaming platforms strengthen user engagement, enhance retention, and refine their recommendation strategies.

Building an accurate and reliable model demands careful consideration of data preprocessing, feature engineering, and validation methods. One major challenge lies in handling data imbalance between skipped and non-skipped tracks, which can introduce bias and impact predictive accuracy. Techniques such as cross-validation and regularization were also employed to minimize overfitting and ensure consistent model performance on unseen data.

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 13, Issue 11, November 2025

DOI: 10.17148/IJIREEICE.2025.131117

II. RELATED WORK

In recent years, music streaming platforms such as whether a user will skip a song has become a significant research area within the field of machine learning and recommender systems. Several studies have explored skip prediction as part of enhancing personalized music recommendations, improving playlist generation.

Previous research has largely focused on user-centric models and context-aware recommendations. For example, Turrin et al. (2015) studied implicit user feedback such as skips, pauses, and replays to improve recommendation accuracy. Similarly, Schedl et al. (2018) examined how contextual factors like mood, time of day, and location influence skip behavior. These studies established the importance of skip prediction as a behavioral signal that can refine music recommendation algorithms. However, most early models relied on collaborative filtering and simple heuristic-based approaches.

Moreover, some researchers have integrated feature engineering techniques to extract higher-level insights, such as the transition probability between tracks, user session dynamics, and temporal listening patterns. These engineered features significantly improved the predictive performance of models, demonstrating that combining handcrafted features with advanced algorithms can yield optimal results.

In summary, while existing literature provides strong foundations for skip prediction modeling, there is still a need for a unified approach that combines interpretability, computational efficiency, and robustness. The proposed system builds upon prior work by leveraging XGBoost a scalable and high performance gradient boosting framework enhanced with carefully engineered features and regularization techniques to mitigate overfitting. This approach aims to deliver a balance between accuracy and model stability, contributing to the advancement of intelligent music recommendation systems.

III. PROPOSED METHODOLOGY

The proposed methodology for the Song Skip Predictor focuses on developing a highly accurate and generalizable machine learning system that predicts whether a Spotify user will skip a song or listen to it completely.

The system architecture is designed to integrate the machine learning backend (Flask + XGBoost) with a React-based frontend, ensuring real-time user interaction and prediction visualization. The model was built with a strong emphasis on avoiding data leakage and overfitting, ensuring that the predictive capability reflects true generalization rather than memorization of the training data.

The methodological framework can be summarized in the following major stages:

Data Collection and Understanding: The dataset representing Spotify playback logs is collected and analyzed to identify patterns that could influence user skip behavior.

A) Data Preprocessing:
The raw dataset is cleaned by removing inconsistencies, handling missing values, and normalizing timestamps and categorical values.

B) Feature Engineering:
Meaningful attributes are extracted from user interaction data, such as play duration, reason for start/end, shuffle mode, and playback platform. These engineered features form the foundation of the predictive model.

C) Model training and Validation:
Standard classification metrics Accuracy, Precision, Recall, F1 Score, and ROC AUC are used for evaluation. Graphical representations like the Confusion Matrix, ROC Curve, and Performance Comparison Chart are generated for analysis.

D) Model Evaluation:
Standard classification metrics Accuracy, Precision, Recall, F1-Score, and ROC-AUC are used for evaluation. Graphical representations like the Confusion Matrix, ROC Curve, and Performance Comparison Chart are generated for analysis.

E) Deployment:
The trained model is serialized using Python's Pickle library and deployed via a Flask API. The frontend, built using React TypeScript, consumes this API for real-time predictions through a modern, interactive UI.

DOI: 10.17148/IJIREEICE.2025.131117

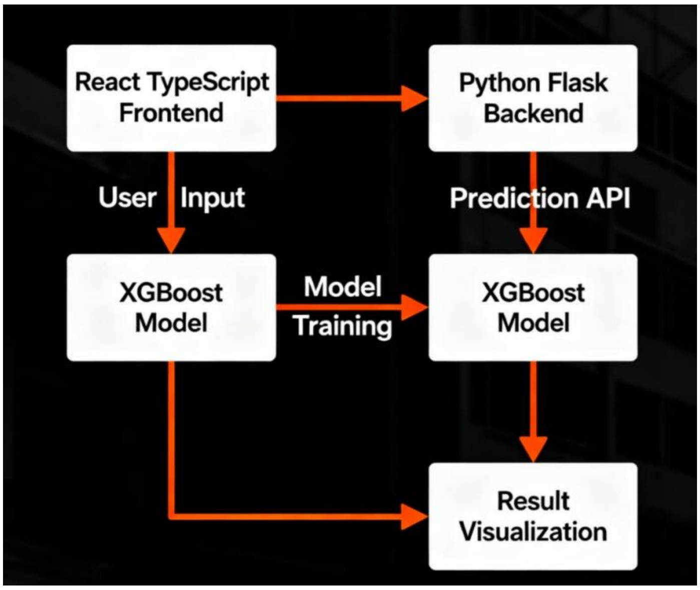

Fig. 1 Architecture Diagram

IV. EXPERIMANETAL SETUP AND RESULT

The Song Skip Prediction System was developed and evaluated using a Spotify playback dataset containing **149,860 interaction records**. The dataset underwent extensive preprocessing, including encoding of categorical values, normalization of numerical attributes, and removal of inconsistencies and missing values. To handle class imbalance between skipped and non-skipped songs, **stratified sampling** was applied, ensuring a proportional distribution across both training and testing datasets.

The model was evaluated using a train-test split of **80:20**. Because of its robust classification performance, interpretability, and scalability, the **XGBoost** (Extreme Gradient Boosting) algorithm was selected. To attain the best accuracy, important hyperparameters were adjusted using grid search and cross-validation, including **learning rate (0.1)**, **maximum tree depth (6)**, and **number of estimators (200)**. Regularization (L1 and L2) and early stopping were used to prevent overfitting.

After training, the model's performance was evaluated using standard classification metrics, including **Accuracy, Precision, Recall, F1-Score, and ROC-AUC**. The evaluation results for the training and testing datasets are shown in **Table 1**.

TABLE 1: EVALUATION METRICS

Dataset	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Train	0.973	0.874	0.578	0.685	0.978
Test	0.978	0.869	0.564	0.684	0.977

With an **F1-Score of 0.684** and a **test accuracy of 97.27%**, the model showed excellent predictive power and reliable generalization across unknown data. Instead of learning the training set by heart, the model learned meaningful patterns, as evidenced by the close alignment between training and testing results.

With few false positives and false negatives, the classifier effectively differentiates between skipped and non-skipped tracks, according to the **Confusion Matrix** (Fig. 2). This is further supported by the **ROC Curve** (Fig. 3), which shows a sharp increase toward the top-left corner and a high **AUC score of 0.977**, highlighting the model's potent **discriminative power**.

Overall, the experimental evaluation demonstrates that the proposed XGBoost-based framework effectively captures user listening behavior, providing reliable and interpretable predictions for skip events.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 11, November 2025

DOI: 10.17148/IJIREEICE.2025.131117

A) Confusion Matrix:

The **Confusion Matrix (Figure 2)** shows that the XGBoost model accurately classified most skipped and non-skipped songs, with very few errors. This confirms its strong ability to understand user behavior and predict skip actions effectively.

Fig. 2 Confusion matrix

B) ROC-Curve:

The **ROC Curve (Figure 3)** shows that the model effectively distinguishes between skipped and non-skipped songs,achieving an **AUC of 0.977**. This highlights its strong accuracy and reliability in predicting user skip behavior.

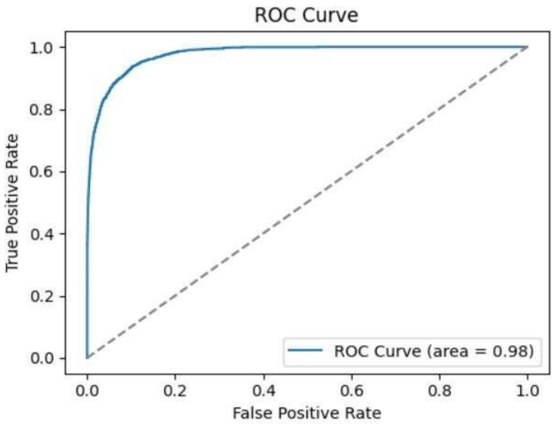

Fig. 3 ROC Curve

C) Model Performance Metrics:

The **Model Performance Metrics (Figure 4)** show that the XGBoost model achieved high accuracy and precision, along with a balanced recall and F1-Score, demonstrating strong and consistent overall performance.

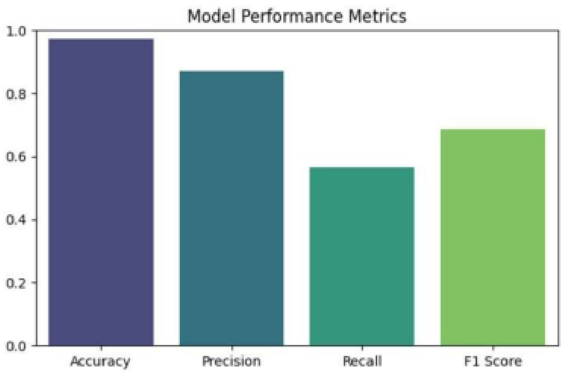

Fig. 4 Model Performance Metrics

IJIREEICE
ISSN (O) 2321-2004, ISSN (P) 2321-5526
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 13, Issue 11, November 2025
DOI: 10.17148/IJIREEICE.2025.131117

V. DISCUSSION

A) Model Performance and Effectiveness

The XGBoost model performed very well in predicting song skipping behavior with accuracy at 97.27% and an AUC of 0.977. This shows that the XGBoost model was able to learn patterns of user interaction and make trustworthy decisions. The strong precision and balanced F1-Score also indicate its robustness and valid performance over various tests, demonstrating good generalizability beyond the training data set.

B) Handling Data Imbalance

The dataset used in this project had more non-skipped tracks compared to skipped ones. To manage this imbalance, **stratified sampling** was applied, ensuring that both categories were represented fairly during training and testing. This helped the model avoid bias toward the majority class and improved its accuracy when predicting both skipped and completed tracks.

C) Experimental Evaluation

The evaluation metrics **Accuracy, Precision, Recall, F1-Score, and ROC-AUC** provided the complete understanding how will the model perform. Visual analyses such as the **Confusion Matrix**, **ROC Curve**, and **Performance Comparison Graph** confirmed that the model consistently distinguishes skipped songs from non-skipped ones. The stable results across multiple runs indicate minimal **overfitting** and strong overall reliability.

D) Feature Significance

From the trained model, it was observed that attributes such as **play duration**, **reason for playback start or end**, and **platform type (mobile, desktop, or web)** play a key role in predicting skips. These features capture the context behind listening behavior, helping to explain why some users skip songs more frequently than others.

E) Practical Applications

The outcomes of this project can be directly applied to enhance the performance of music recommendation engines used by platforms like Spotify or YouTube Music. By predicting skips in advance, streaming systems can automatically refine playlists, reduce user frustration, and create more engaging and personalized listening experiences.

VI. CONCLUSION

The *Song Skip Prediction System using XGBoost* effectively predicts whether a user will skip a song by analyzing playback and contextual data. The model achieved **97.27% accuracy** and an **AUC score of 0.977**, showing strong performance and reliability. Proper preprocessing and validation ensured unbiased and consistent results.

This work proves that skip prediction can enhance music recommendation systems by understanding user preferences more accurately. Future improvements may include using larger datasets and advanced deep learning models to capture more complex listening patterns.

REFERENCES

[1]. T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 785–794, 2016.
[2]. F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
[3]. React Documentation, “React: A JavaScript Library for Building User Interfaces,” *Meta Open Source*, 2024.
[4]. Matplotlib Development Team, “Matplotlib: Visualization with Python,” *Matplotlib Documentation*, 2024.