Music Genre Classification & Recommendation

Siddharth M. Haveliwala
shaveliw@buffalo.edu
University at Buffalo

Manasi J. Dudhane
manasija@buffalo.edu
University at Buffalo

Siddarth G. Gundrathi
sgundrat@buffalo.edu
University at Buffalo

Abstract— **The rapid evolution of music streaming platforms has necessitated effective music recommendation systems catering to individual preferences and moods. While existing research explores collaborative filtering, content-based filtering, and hybrid techniques, this paper addresses gaps by proposing a predictive model leveraging user browsing history and Spotify's extensive dataset for personalized playlist generation. The study integrates insights from related works and explores correlations in music features. Data analysis includes user browsing history and top hits from 2000 to 2022, employing visualization and pre-processing. The research methodology involves data loading, exploration, visualization, correlation analysis, and feature selection. Model selection considers Decision Tree, XGBoost, Ada Boost, Cat Boost, and Random Forest, evaluating their effectiveness in music genre prediction. The paper concludes with observations on dataset bias, model performance, and outlines future directions for advanced recommendation systems.**

Keywords — music recommendation, predictive model, user browsing history, Spotify dataset, data visualization, feature selection, machine learning, decision tree, XGBoost, AdaBoost, CatBoost, Random Forest, Cross-validation.

I. Introduction

With the rapid evolution of music streaming platforms, the need for effective music recommendation systems has become increasingly pronounced. While collaborative filtering, content-based filtering, and hybrid techniques have been explored in existing literature, there remains a gap in seamlessly combining user-specific preferences with a diverse array of music features. This study aims to bridge these gaps by proposing a predictive model that integrates user browsing history and Spotify's extensive music dataset. Unlike prior approaches, our focus is on refining the accuracy and specificity of personalized music recommendations. The literature review underscores the inadequacies of existing studies in incorporating user browsing history, and our research methodology involves comprehensive data analysis, visualization techniques, and a thoughtful model selection process. In this paper, we not only present a novel model but also provide insights into correlations between music features and genres, contributing to the advancement of music recommendation systems.

II. Literature Review

The rapid evolution of music streaming platforms has led to an increased need for effective music recommendation systems that cater to individual preferences and moods. Existing research has explored various methodologies, such as collaborative filtering, content-based filtering, and hybrid recommendation techniques, to address the challenges of playlist generation. Motivated by the limitations observed in current approaches, this study aims to contribute to the field by developing a predictive model for personalized music playlist generation. Our focus is on leveraging user browsing history and Spotify's extensive music dataset to enhance the accuracy and specificity of recommendations based on user interests, moods, and genres.

While the literature provides valuable insights into different aspects of music recommendation systems, there exists a research gap in terms of combining user-specific preferences with diverse music features. Most existing studies focus on collaborative filtering, content-based filtering, or hybrid approaches without giving sufficient attention to the incorporation of user browsing history. Additionally, there is a need to explore the effectiveness of predictive modeling in tailoring playlists to specific user interests and moods, considering the nuances of historical preferences. This research addresses these gaps by proposing a model that integrates user browsing data and Spotify's extensive music history to generate personalized playlists.

Chodos et al. (2019**)** explored the characteristics of musical features accessible through the Spotify API, standardizing input music information. Their study highlighted the efficacy and constraints of collaborative filtering, content-based filtering, and hybrid recommendation techniques. Notably, this research laid the groundwork for playlist depiction using standardized arrays and geometric clustering, focusing on 'genres.' However, the recommendations were not influenced by a user's listening or browsing history.

Dhivya and Mohandas (2023) utilized convolutional neural networks and k-nearest neighbor classifiers for music instrument recognition, focusing on monophonic audio files. While their approach contributes to the field of feature extraction, it is limited to monophonic music and does not consider the polyphonic nature of most music genres.

Lee et al.(2015) proposed a recommendation system based on real-time user brainwaves and genre feature classification. Their system utilized Mel-frequency cepstral coefficients, Decorrelated filter bank, and Octave-based spectral contrast for feature extraction. The study, however, did not incorporate user browsing history in generating recommendations.

Lu and Tintarev (2018) addressed the research gap between diversity-based and personality-based recommender systems. Their algorithm adjusted diversity degrees in music recommendations based on user personalities. The study utilized Spotify Web API and considered specific attributes like release times, number of artists, genres, and audio features. However, it did not focus on personalized recommendations driven by user browsing history.

Bertram et al. used knowledge graph embeddings to enhance music recommendations through the EARS system. The study emphasized the importance of domain-specific knowledge but did not delve into the influence of user browsing history on recommendations.

Pérez-Marcos and López Batista collaborative filtering-based study proposed a recommendation system solely reliant on Spotify's API, generating recommendations based on frequently played songs. The research showcased the effectiveness of a Spotify-specific recommender system without external data dependencies.

Helmholz, Meyer, and Robra-Bissantz (2019) conducted a music research project focusing on a novel music recommendation application centered around emotions and intuitions. The study used Spotify and its accessible API to generate playlists tailored to users' selected genres and emotional preferences. The research emphasized recognizing users' desire to alter their emotional states rather than solely seeking support in their current moods.

Rumiantcev and Khriyenko (2020) developed an emotion-based music recommendation system using advanced technology to suggest music that supports mental and physical health. The system created playlists based on people's mood and activity for specific groups, aiming to solve the problem of choosing music for diverse emotional experiences.

Chen and Chen (2001) developed a Music Recommendation System (MRS) that provided personalized music suggestions by analyzing MIDI format music items. Their approach focused on unique features extracted from polyphonic music objects, achieving an 83% accuracy rate in analyzing MIDI files.

While the existing literature provides valuable insights into various aspects of music recommendation systems, there is a noticeable research gap in incorporating user browsing history for personalized playlist generation. This study aims to address this gap by developing a predictive model that leverages both user browsing history and Spotify's extensive music dataset to enhance the accuracy and specificity of music recommendations. The proposed model aims to provide a comprehensive solution that considers individual user preferences, moods, and genres, contributing to the advancement of personalized music playlist generation.

III. DATA

In this exploration of data analyses, the sources from which data is collected are introduced. The section continues with visualization, pre-processing that has taken place, and introducing the final dataset that is incorporated in this study. It is important to note that the dataset for user's music browsing is not yet available with us and hence not explored in this analysis.

A. *Data Source and Description*

The data used in this study is composed of two general types of information, user's music browsing history and top hit songs for the past few years. The scope of the user's music browsing history incorporates data for the past year. While the top hit songs data for this study covers top 100 songs of the year extending from 2000 to 2022. Following, detailed explanation is provided.

1. Data Description

1.1. Top Hit Songs from Spotify (2000-2022)

The dataset contains 2300 samples for the period between 2010 – 2022.

1.2. Music Features

In this study, we have identified and learned the patterns and relationships between the feature variables like *danceability, energy, loudness, valence, instrumentalness etc.* Analysis of the behavior of these features help us to classify the specific music genre based on the values.

1.3. Response

The response is a list of songs belonging to a specific genre (class) based on the user request. This data can be found by observing the feature patterns, defining clusters and classification based on the important music features. By applying feature engineering, clustering based on K-means method, evaluation of clustering we aim to curate a playlist by labelling genres from user's music browsing data (90%) and song suggestions from top hit songs data (10%).

B. *Pre-Processing*

2.1. Identifying missing values

Missing values in the dataset are identified and removed to reduce bias.

2.2. Removal of unnecessary features.

Removed features like *playlist_url, track_id, artist_id* and *artist_genres* as these features do not provide any information on the response.

2.3. Standardization of data

The top hit songs data already have the music features values standardized to ease the model development and testing.

C. *Visualization*

3.1. Histogram Plots

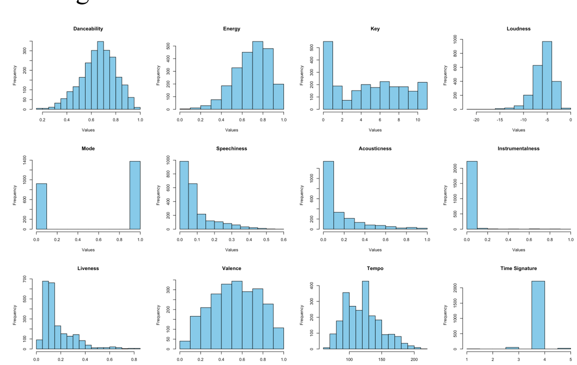

Figure 3.1: Histogram plots for music features

From figure 3.1 it is evident that the *danceability* and *valence* histograms show a pattern that resembles a normal distribution, suggesting that these musical characteristics are distributed fairly across the songs that were examined. Top chart music is known for their higher energy, as proven by a significant rise in *energy* levels, especially in songs scoring over 70%. In terms of mode feature, the dataset's dominance of value 1 over 0 is highlighted by its wide distribution and mean of 0.5985. For more than 70% of the songs, *speechiness* primarily lies between 0 and 0.1, indicating a tendency towards instrumental richness. In the case of *instrumentalness*, a surprising finding is that more than 95% of songs have values less than 0.1, indicating that most popular songs have lyrics.

With a mean of 0.5351, feature *valence* is in line with expectation and suggests that the top charts tend to have enjoyable tracks. Moreover, about 95% of songs have a *time signature* close to 4, which is consistent with industry norms and general adherence to the conventional 4-beats-per-bar structure. These findings give insight on the repeating musical patterns among the songs that are at the top of the charts, which gives the study vital context.

3.2. Boxplots

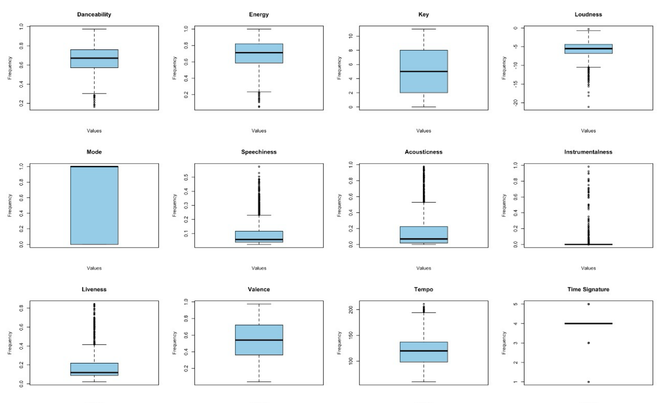

Figure 3.2: Box plots for music features

From figure 3.2 it is observed that the box plot for *danceability* shows that most songs have moderate *danceability*, with the median falling at 0.6710. However, it also reveals some outliers with low *danceability*, suggesting a diversity of dance-friendly tracks. Most songs exhibit a balanced range of *energy* as seen in the boxplot, which is evident by the median at 0.7120. However, there are a few outliers with low *energy*, implying the presence of slow-paced tracks in the dataset.

The boxplot for *key* follows approximately uniform distribution in the dataset with the range of values from -1 to 11. It indicates the most common key, with a median value of 5 with no outliers. Distribution of *valence* suggests that all types of tracks are covered in the dataset, clustered around the median value of 0.5400 and mean 0.5351. The plot for *'Loudness'* indicates a distribution with a clear negative skew and most of the songs tend to have higher *loudness* levels. It can also be observed from the plot that there are multiple outliers with low *loudness*. *Mode* boxplot highlights the bimodal nature of the feature, with only two peaks in the distribution at 0 and 1 with *mode* as 1. For *'Speechiness*,' the box plot displays a positively skewed distribution with a median of 0.0568. The data indicates that most of the songs have low *speechiness* values, although there are multiple outliers explaining the variation of music in the dataset. *Acousticness* boxplot suggests that most songs in the dataset have lower *acousticness* explaining the high rated tracks, but also has a few outliers featuring higher acoustic characteristics.

For the *Instrumentalness* feature, the values above 0.4 are less on vocals hence most of the data points hover at zero with some outliers explaining the variation in tracks. It is quite evident that most songs have low *liveness* as they are recorded, with few outliers valued above 0.4. The *'Tempo'* box plot reveals a relatively symmetrical distribution with a median tempo of 120. The range of Time signature is 3 to 7. *Time Signature* boxplot highlights expected behavior of the feature with value of 4, within the original range from 3 to 7. However, a few outliers do exist. From figure 1 it is evident that the *danceability* and *valence* histograms show a pattern that resembles a normal distribution, suggesting that these musical characteristics are distributed fairly across the songs that were examined. Top chart music is known for their higher energy, as proven by a significant rise in *energy* levels, especially in songs scoring over 70%. In terms of *mode* feature, the dataset's dominance of value 1 over 0 is highlighted by its wide distribution and mean of 0.5985. For more than 70% of the songs, *speechiness* primarily lies between 0 and 0.1, indicating a tendency towards instrumental richness. In the case of *instrumentalness*, a surprising finding is that more than 95% of songs have values less than 0.1, indicating that most popular songs have lyrics.

With a mean of 0.5351, feature *valence* is in line with expectation and suggests that the top charts tend to have enjoyable tracks. Moreover, about 95% of songs have a *time signature* close to 4, which is consistent with industry norms and general adherence to the conventional 4-beats-per-bar structure. These findings give insight on the repeating musical patterns among the songs that are at the top of the charts, which gives the study vital context.

3.3. Correlation Plots

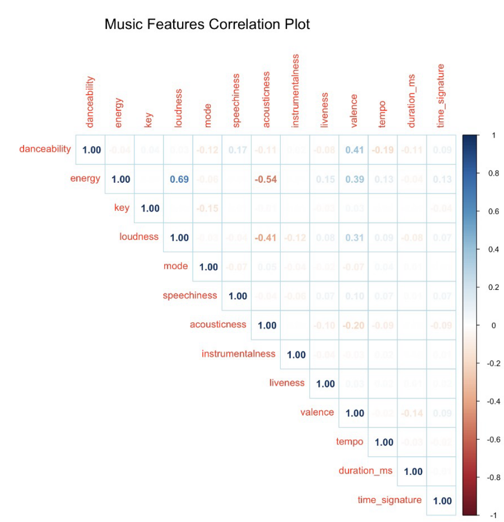

Figure 3.3: Correlation plot

Among the musical features, there exists numerous significant relationships in our study as shown in figure 3. 3 Notably, there is a significant positive association (0.69) between *loudness* and *energy*, meaning that louder music is typically associated with higher energy levels. On the other hand, there is a moderate negative correlation (-0.54) between *acousticness* and *energy*, which indicates that songs that are acoustic typically have lower energy levels. Louder songs are also generally less acoustic, as seen by the negative correlation (-0.41) between *loudness* and *acousticness*. *Danceability* and *valence* shows an evident positive correlation (0.41) in terms of sentimental tone, indicating that songs with higher *danceability* typically have happier moods.

Songs with higher energy levels typically convey a more positive emotional character, according to the slight positive correlation between *energy* and *valence* (0.39). These correlations provide light on the complex interactions between different musical attributes and offer insightful information about the links between various musical characteristics.

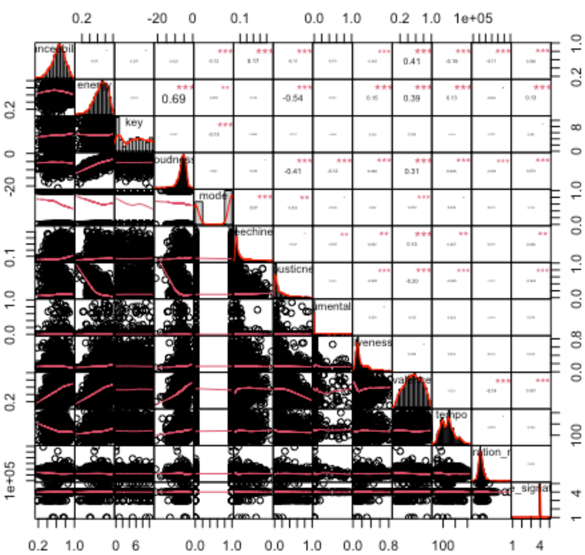

Figure 3.4: Correlation and distribution plot

Figure 3.4 shows the pairwise correlation matrix with histogram of all the variables collected for this study. Note the red lines, indicating the correlation smooth line and the distribution in the histogram charts.

As observed, there are multiple correlations between multiple independent features with a few strong correlations which can also be viewed in the plot. The *danceability* and *valence* explain the variance of the data reasonably well. It is also evident that features *duration_ms* and *time_signature* have no correlation with any of the other features, which is expected as these are not core response defining variables. Between these two, we have decided to pick only *duration_ms*, based on the observations noted here. To reduce redundancy in the model because of strong positive correlation, we can choose one of the two features *loudness* and *energy.* For the analysis conducted here, number of categorical variables which are less significant for modelling are omitted.

3.4. Scatter Plots

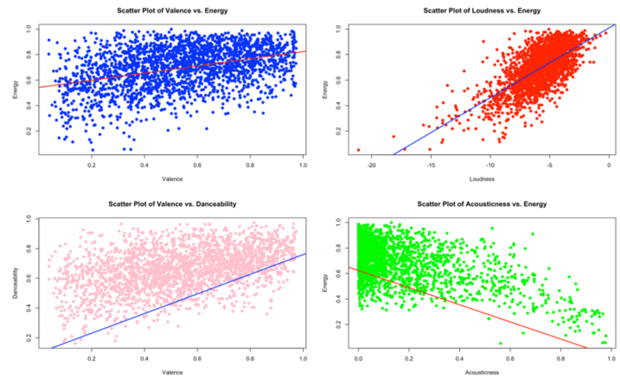

Figure 3.5: Correlation and distribution plot

The scatter plots in our exploratory data analysis give insight into relationships between important musical features. The plots of *valence* versus *energy*, *loudness* versus *energy*, and *valence* versus *danceability*, in particular show positive associations. Regression lines suggest that songs with higher levels of *energy* and *danceability* tend to have higher *valence*. The *Acousticness* versus *Energy* plot, on the other hand, indicates a negative association, suggesting that songs with higher *energy* levels are typically less acoustic. This finding offers a useful classification insight.

IV. RESEARCH METHODOLOGY

4.1. Data Loading and Preprocessing:

Our analysis of the dataset commenced with the loading of the data, consisting of instances and features essential for understanding musical characteristics. The dataset was meticulously preprocessed to ensure optimal compatibility with machine learning models. This preprocessing phase involved cleaning the data, transforming it into a suitable format, and encoding it to focus on numeric attributes. Addressing inconsistencies and handling missing values were integral steps to maintain data quality and reliability. The refined dataset, emphasizing numeric attributes, sets the stage for subsequent model training and evaluation in our analysis of music genres.

Aligned with the numerical-centric nature of machine learning algorithms, a key aspect of our preprocessing involved selecting relevant features. Non-numeric columns were omitted to streamline the dataset, concentrating on numeric attributes crucial for machine learning models. Simultaneously, missing values were addressed to maintain data integrity, ensuring that our models are trained and evaluated on complete and reliable data. This precise preprocessing contributes to the robustness of our analysis, preparing the dataset for in-depth exploration of features and the development of machine learning models for predicting music genres.

In the Spotify dataset, various features provide insights into the musical characteristics of tracks. The definitions and values for these features are detailed in the accompanying table:

Feature	Description	Example
Acousticness	A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.	0.00242
Danceability	Describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity.	0.585
Energy	A perceptual measure of intensity and activity, ranging from 0.0 to 1.0. Typically, energetic tracks feel fast, loud, and noisy.	0.842
Instrumentalness	Predicts whether a track contains no vocals. The closer the value is to 1.0, the greater likelihood the track contains no vocal content.	0.00686
Key	The key the track is in, mapped to pitches using standard Pitch Class notation. Range: -1 to 11. If no key was detected, the value is -1.	9
Liveness	Detects the presence of an audience in the recording. Higher values represent an increased probability that the track was performed live.	0.0866
Loudness	The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and typically range between -60 and 0 dB.	-5.883
Mode	Indicates the modality (major or minor) of a track. Major is represented by 1, minor by 0.	0
Speechiness	Detects the presence of spoken words in a track. Values closer to 1.0 describe tracks that are probably made entirely of spoken words.	0.0556
Tempo	The overall estimated tempo of a track in beats per minute (BPM).	118.211
Valence	A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive, while low valence sounds more negative.	0.428

Table 4.1 – Audio Features Definitions

4.2. Data Exploration and Visualization:

Upon conducting a comprehensive exploration of the Spotify dataset, our analysis revealed a predominant distribution of genres. The dataset predominantly comprises Pop, followed by Hip Hop and Rock genres. To streamline our analysis and focus on specific genres, we strategically selected Pop, Hip Hop, and Rock as our variables of interest. This targeted approach ensures a more nuanced examination of the dataset, allowing us to delve deeper into the characteristics and patterns associated with these prominent music genres.

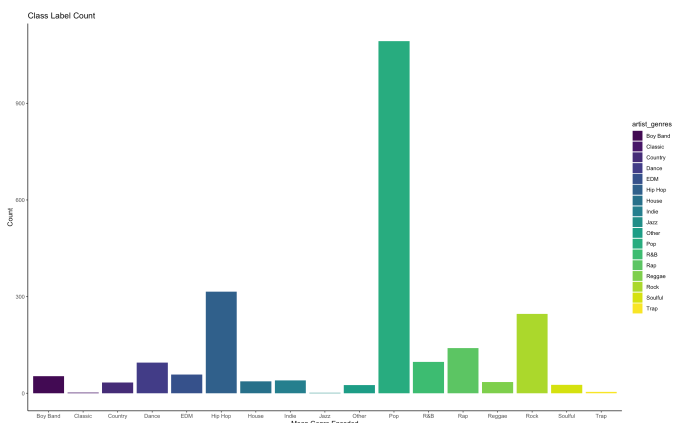

Figure 4.2: Distribution of genres across dataset

4.3. Encoded Genres:

The dataset encompasses a diverse array of genres, including but not limited to Indie, Rock, Country, Boy Band, R&B, Hip Hop, Pop, House, Dance, and others. In our analysis, to facilitate a more detailed examination of the selected genres, we implemented an encoding scheme. Specifically, we encoded Pop, Hip Hop, and Rock as numeric labels 6,11 and 15, respectively, while assigning the label 1 to all other genres. This encoding allows for a simplified representation of the classification task, focusing on the primary genres of interest. To visually represent the distribution of these encoded genres, we created a graph that further enhances our understanding of the dataset's genre composition. This visualization serves as a foundation for our subsequent analyses, providing a clear and concise overview of the selected genres' prevalence in the dataset.

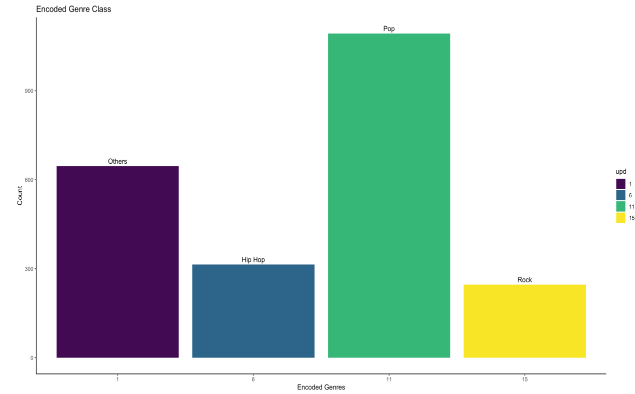

Figure 4.3: Encoded genres distribution

In this context, 'Genre' serves as our response variable—the variable we aim to predict based on the encoded labels. To visually represent the distribution of these encoded genres, we created a graph that further enhances our understanding of the dataset's genre composition. This visualization, centered around 'Genre' as our response variable, serves as a foundation for our subsequent analyses, providing a clear and concise overview of the selected genres' prevalence in the dataset.

4.4. Correlation Analysis and Feature Selection:

In predictive modeling, feature selection is crucial since it determines machine learning algorithms' overall effectiveness, interpretability, and performance. This work advances understanding of ideal procedures to build efficient predictive models for various applications by understanding the relation between feature selection and model selection.

Subsequently, we calculated correlation matrix of the numeric features. This brings to the observation that there is high correlation between specific predictors (‘Danceability’, ‘Acousticness’, ‘Speechiness’, ‘Instrumentalness’ and ‘Energy’) and the target variable ‘Genre’ highlights potential predictive power. This helped in choosing the most relevant features, reducing dimensionality, and potentially improving model performance.

Figure 4.4: Correlation matrix of top 5 features

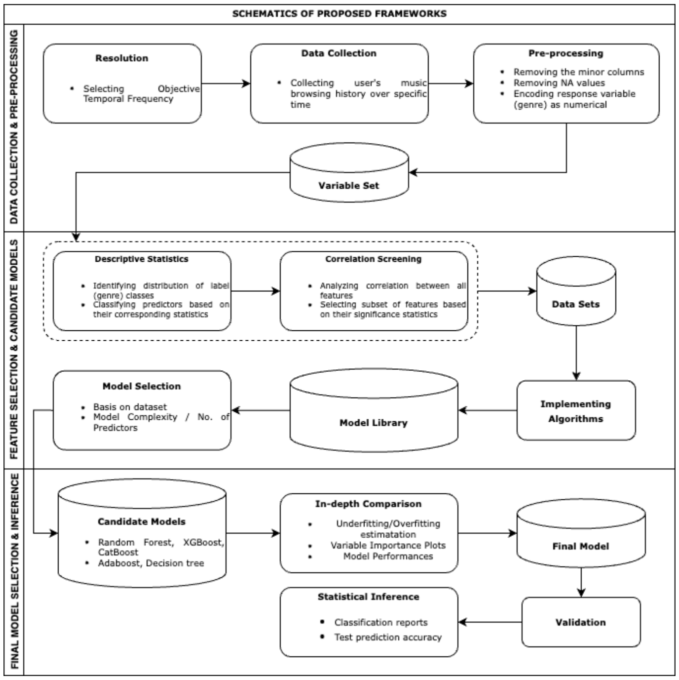

Figure 4.5: Schematics of the proposed framework

V. MODEL SELECTION

Model selection in machine learning is vital for choosing the most suitable algorithm based on the specific problem at hand. It optimizes performance metrics, addresses the bias-variance trade-off, ensures computational efficiency, and balances interpretability with predictive power.

We have partitioned the dataset into 75% and 25% for training and validation sets respectively utilizing 5-fold cross validation method to ensure robust model evaluation, optimize parameters, and obtain reliable estimates of a model's performance on diverse subsets of the data.

5.1. Decision Tree:

A decision tree model is a supervised machine learning algorithm that recursively partitions data into subsets based on key features to create a tree-like structure for classification or regression tasks.

Our implementation of Decision Tree model generates and evaluates a decision tree model for a classification task. The *rpart* function builds the model using the training data, and predictions are made on a test set. The accuracy of the model is then assessed using a confusion matrix, providing a quantitative measure of its performance on the test data. In our evaluation, the decision tree model demonstrated commendable performance with an overall accuracy of 70%. Specifically, it successfully captured 67.4%, 70.4%, and 69.6% of instances for Classes 2, 3, and 4, respectively. These recall values provide insights into the model's effectiveness in correctly identifying instances for each class, contributing to a comprehensive understanding of its classification performance.

5.2. Extreme Gradient Boosting:

Extreme Gradient Boosting (XGBoost) is an ensemble learning algorithm that sequentially combines outputs from weak learners, typically decision trees, to create a robust predictive model. It features regularization techniques to control complexity and is widely utilized for regression and classification tasks due to its high predictive accuracy.

Subsequently, we have implemented multi-class classification using XGBoosting technique. Firstly, the target variable is numerically encoded for training and testing datasets. XGBoost parameters, including the objective function, tree depth, and regularization terms, were specified. The model was trained using the training data. Predictions were made on the test data, and a confusion matrix is generated. XGBoost model accurately classified approximately 68.11% of the instances in the test set. The model achieved recall rates of approximately 69.23%, 69.44%, and 52.94% for Classes 0, 1, and 2, respectively, indicating its effectiveness in correctly identifying instances for each class.

5.3. Ada Boosting:

AdaBoost is an ensemble learning method that sequentially trains weak classifiers, assigning higher weights to misclassified instances to create a robust predictive model. This method applied AdaBoost, using the bagging function from the *ipred* package to a multi-class classification problem. It involved preprocessing the data, training the AdaBoost model with specified features, making predictions, and evaluating accuracy through a confusion matrix.

This method provided intriguing recall results for the class 2 with 84.09% and the model achieved an accuracy of approximately 69.32%, indicating the proportion of correctly predicted instances out of the total instances evaluated.

5.4. Cat Boosting:

CatBoost is a gradient boosting algorithm that excels in handling categorical features, automatically addressing feature importance and overfitting, making it robust and efficient for machine learning tasks.

The confusion matrix for the classification model reveals varied performance across three classes (2, 3, and 4). For Class 2, the model showed a recall of 0.31, indicating it correctly identified 31% of Class 2 instances, which is relatively low. In contrast, the model performed significantly better for Class 3, with a high recall of 0.92, successfully identifying 92% of the instances. However, the performance dropped again for Class 4, with a recall of 0.43, indicating the model correctly identified 43% of Class 4 instances. This disparity in recall values suggests the model is most effective in recognizing Class 3 but struggles with Classes 2 and 4, indicating potential areas for model improvement or data imbalance issues.

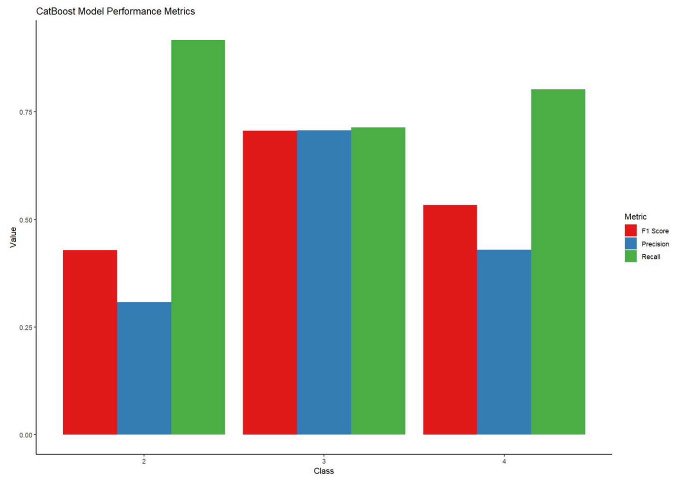

Figure 5.4: CatBoost Performance

5.5. Random Forest:

Random Forest is a method that builds multiple decision trees using ensemble methods and by using random subsets of the training data and features. It aggregates the predictions of these trees to improve model accuracy and generalization, making it robust against overfitting.

This method built a Random Forest classifier with reduced tree depth, performed 5-fold cross-validation, made predictions on a test set, and evaluated model accuracy using a confusion matrix. The Random Forest model achieved an accuracy of approximately 67.87%. The model resulted into recall rates of approximately 38.29%, 84.47%, and 39.29% for Classes 2, 3, and 4, respectively. These values signify the proportion of instances in each class that were accurately captured by the model out of the total instances of that class. The higher the recall, the more effectively the model identifies instances for the corresponding class.

VI. Conclusion

In conclusion, our model selection process encompassed rigorous evaluations of various machine learning algorithms, including Decision Tree, Extreme Gradient Boosting (XGBoost), Ada Boosting, Cat Boosting, and Random Forest. Employing 5-fold cross-validation with a dataset split of 75% for training and 25% for validation, each algorithm underwent thorough scrutiny.

Cat Boosting, a gradient boosting algorithm tailored for categorical features, emerged as the standout model in our evaluation. Despite varied recall rates across classes, with 31% for Class 2 and 43% for Class 4, it excelled with an impressive 92% recall for Class 3. The overall robustness and efficiency of Cat Boosting, particularly in handling categorical features, position it as a promising choice for our multi-class classification task.

We've identified bias in our dataset, particularly in the accuracy of Pop songs due to the large quantity of samples from this genre. Recognizing this bias is crucial as it falls below industry standards. Understanding how bias affects overall accuracy underscores the need for a more targeted strategy. We've observed significant improvements in model performance by focusing on key genres like Pop, Hip-Hop, Rock, and Other. By addressing bias within these specific genres, our model reaches an acceptable level of accuracy, effectively resolving the problem at hand.

VII. Future scope

Further into the future we plan to implement cutting-edge algorithms that leverage larger datasets to enhance accuracy and diversity to build advanced recommendation systems. This evolution involves venturing into new territories by not only including niche genres but also exploring sub-genres to cater to an even broader range of user preferences. To create a more connected and enjoyable listening experience, forming genre-specific user communities becomes pivotal. These communities serve as hubs for personalized music suggestions, fostering a deeper connection between users and their music preferences, ultimately shaping a more tailored and enriching musical journey.

References

[1] Chodos AT. Solving and dissolving musical affection: A critical study of spotify and automated music recommendation in the 21st century. [Order No. 27545374]. University of California, San Diego; 2019.

[2] Dhivya, S., Mohandas, P. (2023). Comparison of Convolutional Neural Networks and K-Nearest Neighbors for Music Instrument Recognition. In: Biswas, A., Wennekes, E., Wieczorkowska, A., Laskar, R.H. (eds) Advances in Speech and Music Technology. Signals and Communication Technology. Springer, Cham. https://doi.org/10.1007/978-3-031-18444-4_9

[3] J. Lee, S. Shin, D. Jang, S. -J. Jang and K. Yoon, "Music recommendation system based on usage history and automatic genre classification," 2015 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 2015, pp. 134-135, doi: 10.1109/ICCE.2015.7066352.

[4] Lu, F., & Tintarev, N. (2018). A Diversity Adjusting Strategy with Personality for Music Recommendation. In P. Brusilovsky, M. de Gemmis, A. Felfernig, P. Lops, J. O'Donovan, G. Semeraro, & M. C. Willemsen (Eds.), Proceedings of the 5th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems (pp. 7-14). (CEUR Workshop Proceedings; Vol. 2225). CEUR.

[5] Niels Bertram, Jurgen Dunkel, Ramon Hermoso. I am all EARS: Using open data and knowledge graph embeddings for music recommendations. Hannover University of Applied Sciences and Arts, Germany, University of Zaragoza, Spain

[6] Javier Pérez-Marcos, Vivian F. López Batista. Recommender System Based on Collaborative Filtering for Spotify's Users. Department of Computing and Automation, University of Salamanca, Plaza de los Caidos s/n, 37008 Salamanca, Spain.

[7] P Helmholz, M Meyer and S. Robra-Bissantz, "Feel the Moosic: Emotion-based Music Selection and Recommendation", Bled eConference, pp. 50, 2019.

[8] M Rumiantcev and O. Khriyenko, "Emotion Based Music Recommendation System", Proceedings of Conference of Open Innovations Association FRUCT 2020.

[9] HC Chen and AL. Chen, "A music recommendation system based on music data grouping and user interests",Proceedings of the tenth international conference on Information and knowledge management, pp. 231-238, 2001 Oct 5.

[10] Li, XinHai. "Using" random forest" for classification and regression." Chinese Journal of Applied Entomology 50.4 (2013): 1190-1197.

[11] https://medium.com/analytics-vidhya/introduction-to-xgboost-algorithm-d2e7fad76b04

[12] https://www.geeksforgeeks.org/catboost-ml/