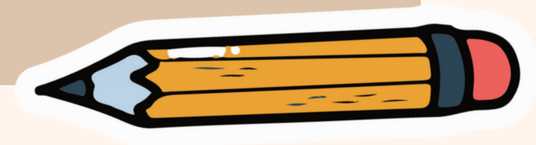
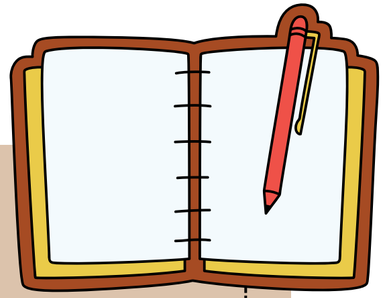
Information Retrieval using Audio file

Problem Statement

The music industry faces a challenge in efficiently categorizing and discovering music genres due to the overwhelming volume and diversity of audio content. Manually assigning genres to vast music libraries is time-consuming and subjective.



Solution



Proposing an automated music genre prediction system utilizing advanced machine learning algorithms and feature extraction. This system aims to enhance efficiency, accuracy, and user experience by streamlining genre categorization, providing real-time predictions, and adapting to the expanding music landscape.

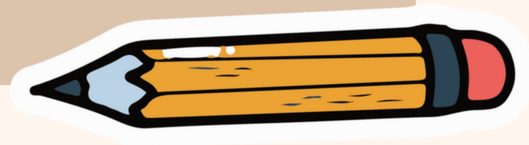
Topic

> Efficiently categorizing and discovering music genres

> Automated system for accurate music genre prediction

Literature Review

Numerous studies have explored the use of spectral features for music genre classification tasks. Researchers have investigated various feature extraction techniques, including spectral centroid, spectral flux, and spectral rolloff, to capture different aspects of the frequency content in music recordings. Machine learning algorithms such as Support Vector Machines (SVM), Random Forests, and Convolutional Neural Networks (CNNs) have been widely employed for classification tasks due to their ability to learn complex patterns from audio features. Existing literature emphasizes the importance of feature selection and model optimization in achieving accurate and robust music genre classification performance.



Possible methodologies

- **Data Preparation:** Load the extracted audio features from the CSV file and preprocess the data (e.g., handle missing values, encode categorical variables, scale features).

- **Feature Selection:** Explore feature selection techniques to identify the most informative features for classification (e.g., Recursive Feature Elimination, mutual information, Principal Component Analysis).

- **Model Training:** Split the dataset into training and testing sets, choose appropriate machine learning models (e.g., XGBoost, SVM, neural networks), and train the models on the training set.

contd...

- **Model Evaluation:** Evaluate the trained models' performance using metrics like accuracy, precision, recall, and F1-score on the testing set. Utilize cross-validation to assess model robustness.

- **Hyperparameter Tuning (Optional):** Fine-tune the models' hyperparameters using techniques like grid search or Bayesian optimization to improve performance further.

- **Final Model Selection and Deployment:** Select the best-performing model based on evaluation metrics and deploy it for inference on new data.

Process

Input

GTZAN Dataset

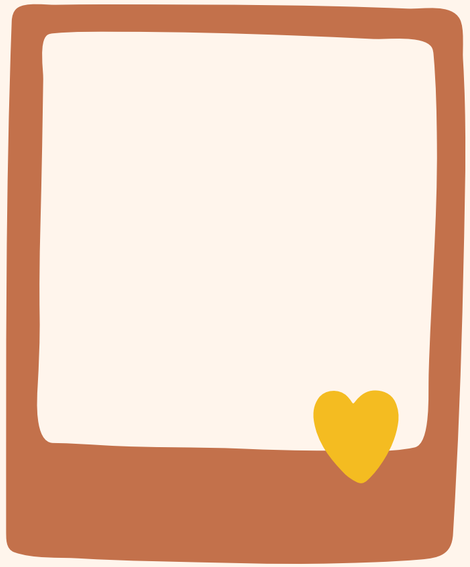
- The GTZAN dataset is a widely used collection of audio files commonly employed for music genre classification tasks.
- The GTZAN dataset consists of 1000 audio tracks, each 30 seconds long, evenly distributed across 10 different music genres. The genres included in the dataset are:,Blues,Classical,Country,Disco,Hip-hop,Jazz,Metal,Pop,Reggae and Rock.

contd...

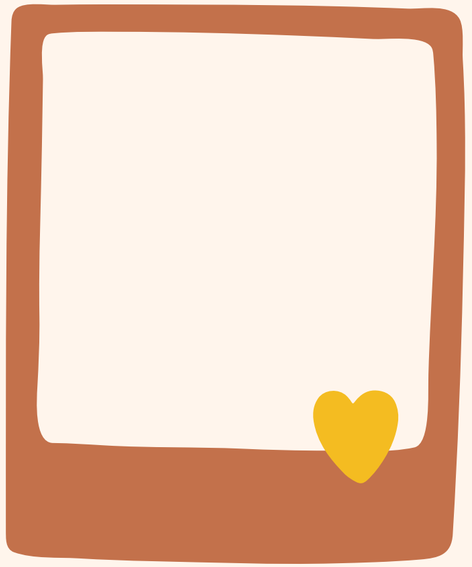
Expected output

- Predicted genre label for each input spectrogram image.
- Each spectrogram image, representing a segment of audio, is classified into one of several predefined music genres.
- The output cobbnsists of a list of spectrogram images along with their corresponding predicted genre labels.
- Predicted genre labels include categories such as Rock, Jazz, Pop, Hip-hop, etc.
- The predicted genre labels provide insights into the genre classification of the input audio segments, aiding in music genre categorization and discovery.

Our Team



Sathvik
PES1UG21CS212



Manasi
PES1UG22CS815

THANK YOU SO MUCH!