**Music Genre Classification Techniques**

**Samarth Gupta**

**Lovely Professional University, Phagwara, India**

**samarth.11911223@lpu.in**

**Abstract:** This is abstract.

**Keywords:** Music Genre classification

1. **Introduction**

As we know Music Industry is rising Every day. The companies like Spotify have contributed a lot to it. But one of the biggest problems they face is Analyse and Categories the data. Classification based on Genre of Music is one of the best categorisations. To do this Automatically, we must analyse the music and then classify into different categories. It can be used to recommend music.

This study explores the application of Machine Learning (ML) algorithms to identify and classify the genre of a given audio file. Also, will put some flash on important features of Audio file.

1. **Dataset**

In this work, we make use of GTZAN Dataset, which is a very popular Dataset for Genre classification. This dataset contains 10 Genre classes (blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae, rock). Each class have 100 Music files in it.

1. **Important Concepts**

**3.1 Audio**

**Sound**: Sequence of vibrations in varying pressure strengths

**Sample Rate**: Number of samples of audio carried per second. (Hz or kHz)

**Fourier Transform**: function that gets a signal in the time domain as input and outputs its decomposition into frequencies

**Spectrogram**: It is a visual representation of the signal strength or loudness or pitch of a signal over time at different frequencies present in a waveform. By this we can see energy levels varying over time. [*spectrogram of one music file*]

**Mel Spectrogram**: The Mel Scale, is the result of some non-linear transformation of the frequency scale. The Mel Spectrogram is a normal Spectrogram, but with a Mel Scale on y axis.

* 1. **Audio Features:**

**Zero Crossing Rate**: the rate at which the signal changes from positive to negative or back.

**Harmonics and perceptrual** : Harmonics are charactersticks that human years can’t distinguish (represents the sound color). Perceptrual understanding shock wave represents the sound rhythm and emotion.

**Tempo BMP (Beats per minute)** : Dynamic programming beat tracker.

**Spectral Centroid**: Indicates where the “center of mass” for a sound is located and is calculated as the weighted mean of the frequencies present in the sound.

**Spectral Rolloff**: It is a measure of the shape of signal. It represents the frequency below which a specified percentage of total spectral energy.

**Mel-Frequency Cepstral Coefficients**: MFCCs of a signal are small set of features( usually about 10-20) which concisely describe the overall shape of a spectral envelope. It models the characteristics of the human voice.

**Chroma Frequencies**: Chroma features are an interesting and powerful representation for music audio in which the entire spectrum is projected onto 12 bins representing the 12 distinct semitones of the musical octave.

1. **Classifiers**

**4.1 Machine Learning Approaches:**

**1 KNN:** This linear classiﬁer is generally used for binary classiﬁcation tasks. For this multi-class classiﬁcation task, the LR is implemented as a one-vs-rest method. That is, 7 different binary classiﬁers are trained. During test time, the class with the highest probability from among the 7 classiﬁers is chosen as the predicted class.

**2. Naive Bayes:** This classifier is based on applying Bayes’ theorem with strong independence assumptions between features.

**3. Decision trees:** A tree shaped (if-else) classifier in which internal nodes represents features, branches represent decision rules and leaf nodes represent outcome.

**4. Random Forest:** A forest contains multiple trees. Similarly random forest algorithms create given number of decision trees and get the best result by means of voting. It reduces the over-fitting by averaging the result.

**5. SVM:** SVMs transform the original input data into a high dimensional space using a kernel trick. The transformed data is linearly separated using a hyperplain

**6. Logistic Regression:**

**4.2 Deep Learning:**

ANN

CNN.

1. **Evaluation**

|  |  |  |
| --- | --- | --- |
| **Classifier** | **Accuracy** | **F-Score** |
| **Machine Learning** | | |
| KNN | 0.68 | 0.7 |
| SVM | 0.66 | 0.67 |
| Logistic Regression | **0.72** | **0.73** |
| Naïve Bayes’ | 0.5 | 0.51 |
| Decision Tree | 0.45 | 0.45 |
| Random forest | 0.63 | 0.64 |
| **Deep Learning** | | |
| ANN | **0.68** | **0.72** |
| CNN | 0.33 | 0.35 |