AUDIO GENRE CLASSIFICATION

Presented By:
Himanshu Gulechha (2348520)
Manashwy Padhi (2348528)
Purusharth Mallik (2348542)

2 MSAIM

Guided By:
Dr. Nizar Banu

Agenda

- Introduction to Audio Genre Classification
- Dataset
- Features
- Algorithms Implemented
- Comparing Results
- Conclusion

Introduction to Audio Genre Classification

- What is it?

  Audio genre classification uses machine learning to automatically categorize music based on its sonic characteristics.

- How does it work?

  1. Feature extraction: Audio signals are converted into numerical data describing rhythm, tempo, timbre, and other musical elements.

  2. Model training: Machine learning algorithms learn patterns from labeled audio samples across various genres.

  3. Prediction: New audio is analyzed, and its genre is predicted based on the trained model.

Dataset - GTZAN

- The GTZAN dataset is the most-used public dataset for evaluation in machine listening research for music genre recognition (MGR).

  The dataset has the following types of data:

- **genres original** - A collection of 10 genres with 100 audio files each.
- **images original** - A visual representation (Mel Spectrograms) for each audio file.
- **2 CSV files** - Containing features of the audio files. One file has features of 30 sec clips, the other has features of 3 sec clips (i.e., the second CSV has 10 times more data than the first)

Features

- **Chroma Frequencies**

It is an effective method for analyzing music where pitch can be adequately categorized (generally into twelve categories) and whose intonation approximates the equitable scale.
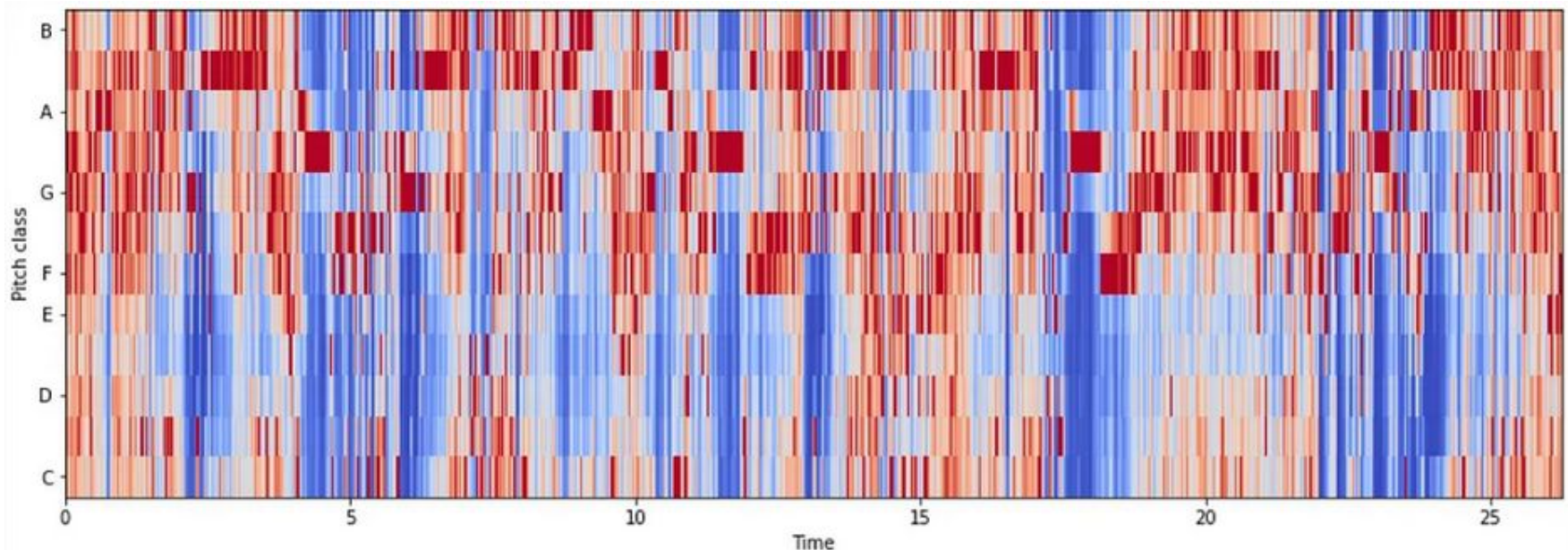
- Mel Frequency Cepstral Coefficients

The mel frequency cepstral coefficients (MFCCs) of an audio signal are a small set of features (usually about 10–20) which describe the overall shape of the spectral envelope.

Mel filter bank
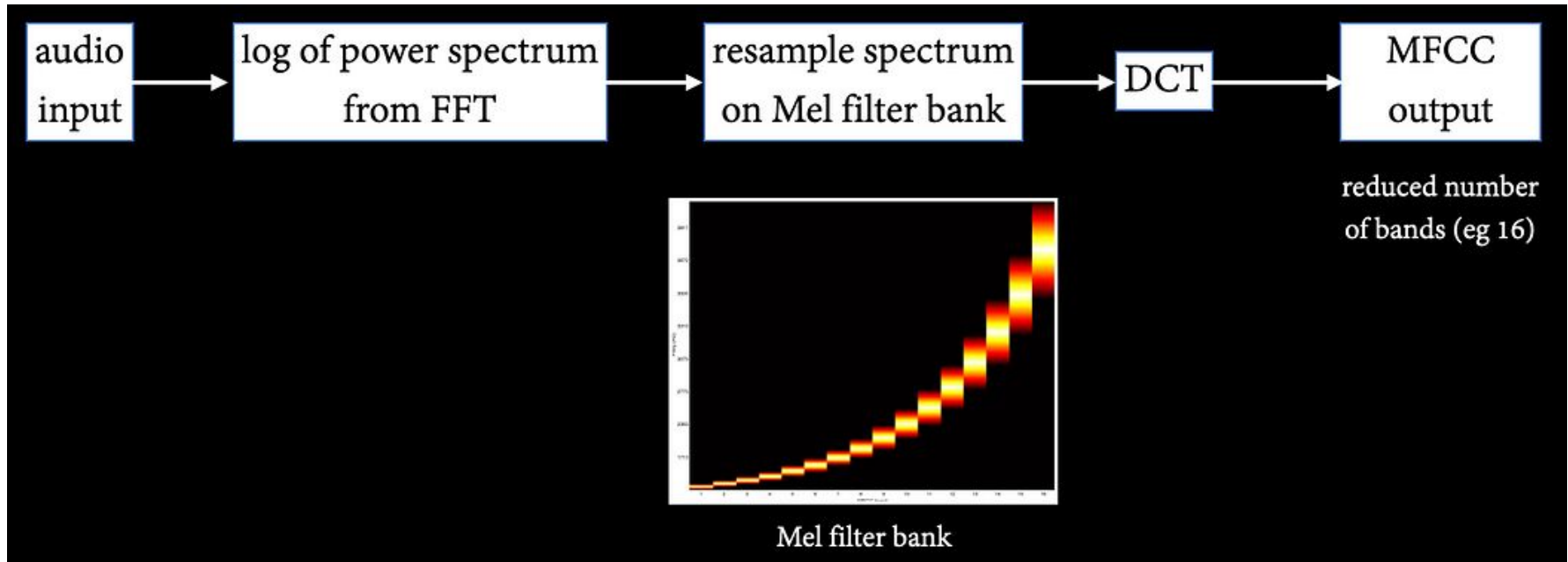
- **Spectral Rolloff**

The spectral rolloff is a parameter that describes the characteristics of a signal's shape. It indicates the frequency at which the high frequencies start to diminish and eventually reach zero.

- **Spectral Bandwidth**

Spectral bandwidth is a measure of the range of frequencies present in a signal.
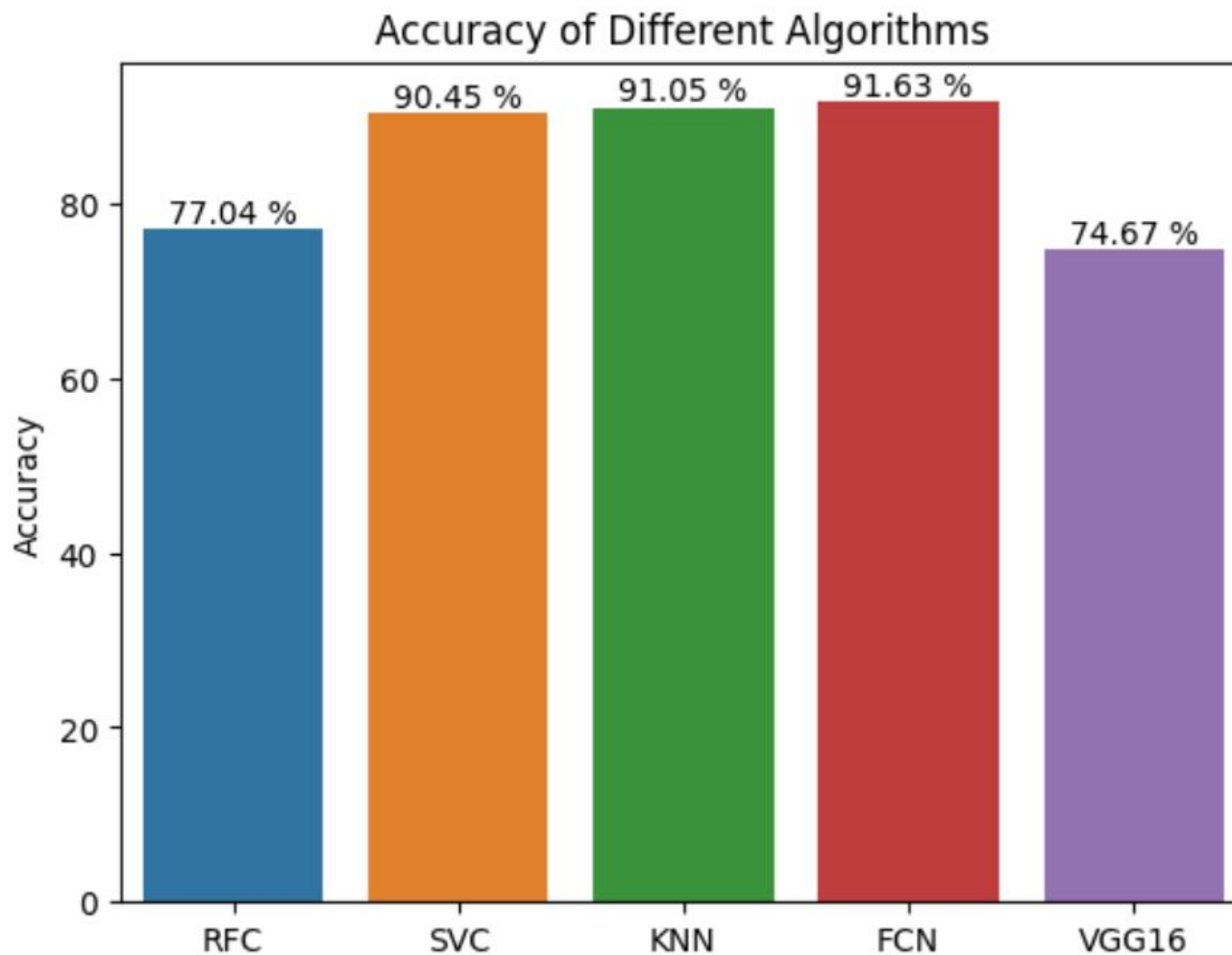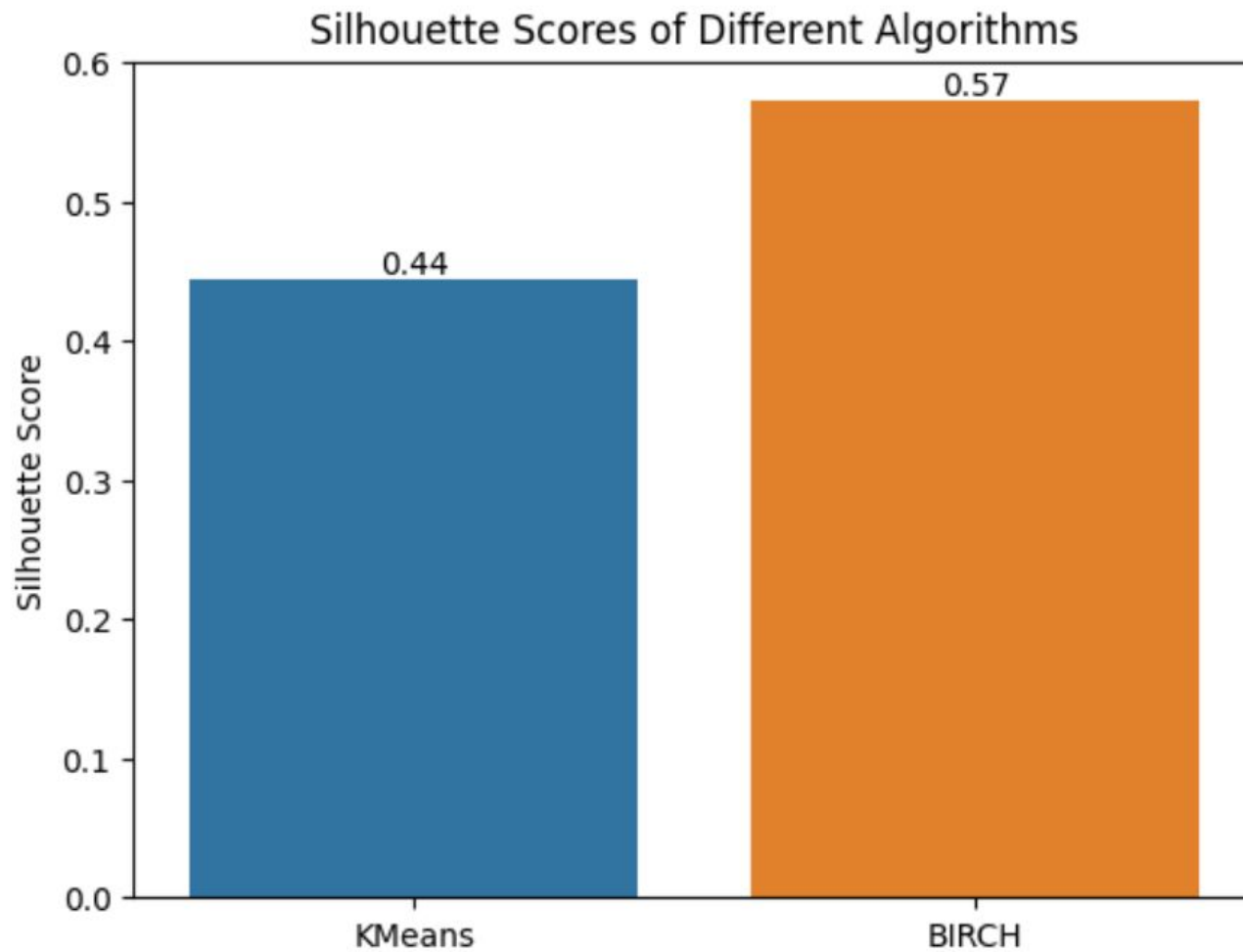
- **Zero Crossing Rate**

The term refers to the speed at which a signal transitions from a positive value to zero, then to a negative value, or from a negative value to zero, and then to a positive value.
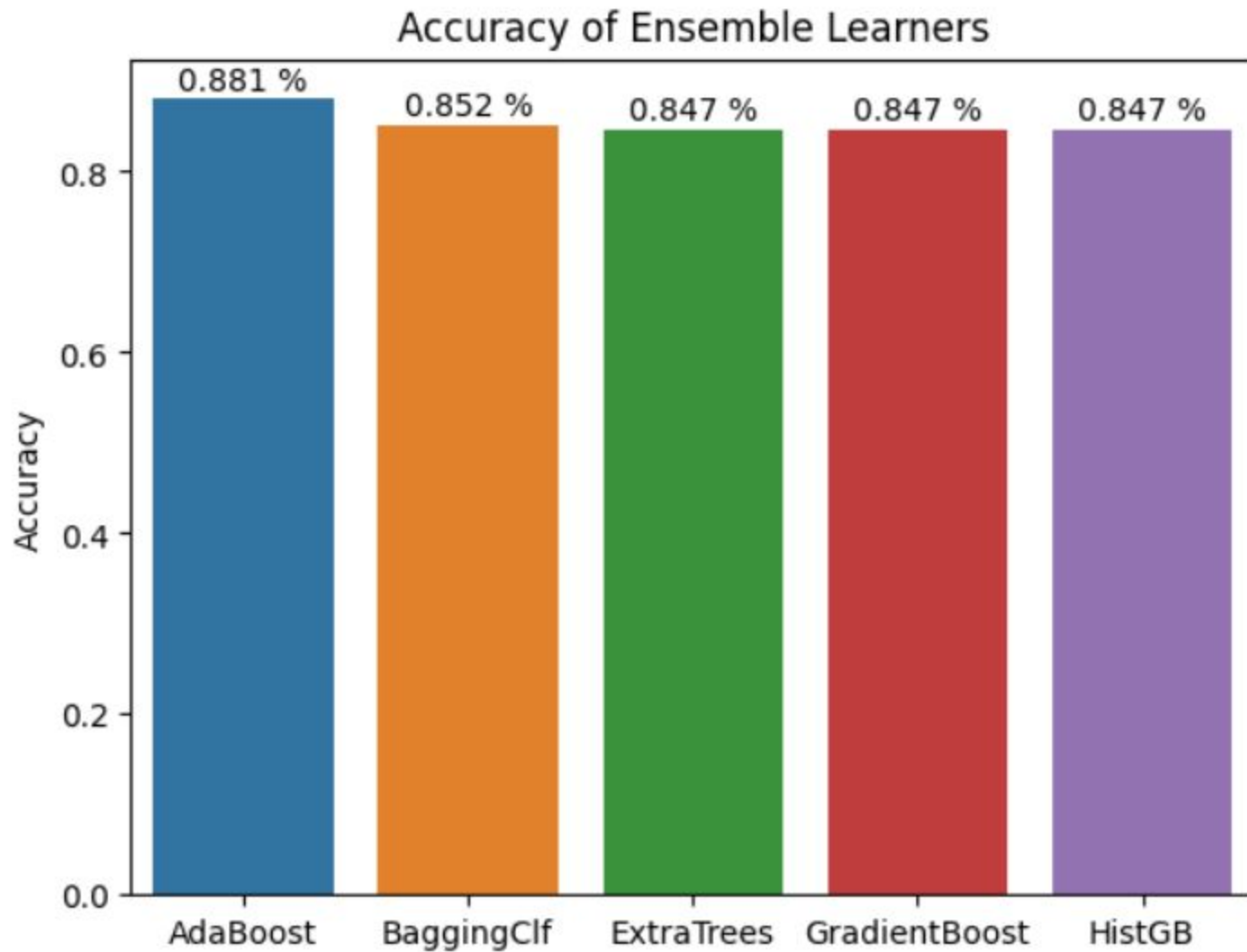
Algorithms Implemented

- **Unsupervised Learning**

- **Dimensionality Reduction**

- **Supervised Learning**

- **Deep Learning**

Comparing Results

Accuracy of Different Algorithms

Silhouette Scores of Different Algorithms

Accuracy of Ensemble Learners

Conclusion

The models that performed the best included Support Vector Classifier, K-Nearest Neighbours, and Fully Connected Network.

Moreover, the models can be fine-tuned if one has access to higher computational power to yield results that are state-of-the-art.

Audio Genre Classification is a sub-problem that occurs while performing various different tasks. With the help of machine learning, we can now classify songs as belonging to a particular genre without the need of human interruption.