Introduction and Objectives

- **Introduction**: Overview of the project aiming to develop a machine learning model for classifying music tracks into different genres.
- **Purpose:** Highlight the objectives and significance of the project, including applications in music recommendation, content organization, market segmentation, and academic research.
- **Objectives:** Mention the primary objectives focusing on model development, feature utilization, and performance comparison.
- **Dataset:** Briefly introduce the "GTZAN Genre Collection" dataset consisting of 1000 audio tracks categorized into 10 genres.

Methodology and Conclusion

- **Methodology:** Describe the approach focusing on implementing and evaluating KNN and Naive Bayes classifiers.
- **Implementation Details:** Provide specifics on implementing each classifier, including parameter tuning and feature space utilization.
- **Evaluation Metrics:** Highlight the performance evaluation metrics, including accuracy, precision, recall, and F1-score.
- **Conclusion:** Summarize the project's objectives, methodologies, and key findings.
- **Future Directions:** Suggest potential avenues for further research and improvement in music genre classification.

DIFFERENCE IN KNN AND NB

	dim	kNN	NB
chroma_cens	84.000000	37.50%	9.99%
chroma_cqt	84.000000	40.03%	1.55%
chroma_stft	84.000000	43.92%	4.20%
mfcc	140.000000	54.99%	41.86%
rmse	7.000000	38.52%	11.78%
spectral_bandwidth	7.000000	45.39%	36.18%
spectral_centroid	7.000000	45.36%	33.31%
spectral_contrast	49.000000	49.55%	39.41%
spectral_rolloff	7.000000	46.25%	28.49%
tonnetz	42.000000	37.31%	22.31%
zcr	7.000000	44.73%	30.39%
mfcc/contrast	189.000000	55.31%	44.03%
mfcc/contrast/chroma	273.000000	53.13%	39.02%
mfcc/contrast/centroid	196.000000	55.23%	43.76%
mfcc/contrast/chroma/centroid	280.000000	53.01%	38.87%
mfcc/contrast/chroma/centroid/tonnetz	322.000000	52.62%	39.06%
mfcc/contrast/chroma/centroid/zcr	287.000000	53.01%	38.90%
all_non-echonest	518.000000	51.77%	9.91%

DIFFERENCE IN KNN AND NB

	kNN	NB
chroma_cens	7.1875	0.0625
chroma_cqt	5.9531	0.0781
chroma_stft	6.3438	0.0312
mfcc	6.2656	0.1094
rmse	0.7500	0.0000
spectral_bandwidth	0.6094	0.0156
spectral_centroid	0.4531	0.0000
spectral_contrast	5.7969	0.0156
spectral_rolloff	0.3750	0.0000
tonnetz	5.6875	0.0000
zcr	0.4844	0.0156
mfcc/contrast	7.3750	0.0938
mfcc/contrast/chroma	8.6406	0.1094
mfcc/contrast/centroid	7.2500	0.0625
mfcc/contrast/chroma/centroid	8.1719	0.1875
mfcc/contrast/chroma/centroid/tonnetz	9.3906	0.1719
mfcc/contrast/chroma/centroid/zcr	8.8906	0.1562
all_non-echonest	13.5625	0.0938

MULTIPLE GENRES

	dim	LR	MLP
mfcc	140.000000	11.23%	12.09%
mfcc/contrast/chroma/centroid/tonnetz	322.000000	12.90%	8.36%
mfcc/contrast/chroma/centroid/zcr	287.000000	13.06%	9.76%

	LR	MLP
mfcc	1.6562	1742.2500
mfcc/contrast/chroma/centroid/tonnetz	6.2812	1437.9844
mfcc/contrast/chroma/centroid/zcr	6.1719	969.3906

BIBLIOGRAPHY


```
@inproceedings{fma_dataset,
  title = {{FMA}: A Dataset for Music Analysis},
  author = {Defferrard, Micha\"el and Benzi, Kirell and
    Vandergheynst, Pierre and Bresson, Xavier},
  booktitle = {18th International Society for Music Information
    Retrieval Conference (ISMIR)},
  year = {2017},
  archiveprefix = {arXiv},
  eprint = {1612.01840},
  url = {https://arxiv.org/abs/1612.01840},
}
```

```
@inproceedings{fma_challenge,
  title = {Learning to Recognize Musical Genre from
    Audio},
  subtitle = {Challenge Overview},
  author = {Defferrard, Micha\"el and Mohanty, Sharada P.
    and Carroll, Sean F. and Salath\'e, Marcel},
  booktitle = {The 2018 Web Conference Companion},
  year = {2018},
  publisher = {ACM Press},
  isbn = {9781450356404},
  doi = {10.1145/3184558.3192310},
  archiveprefix = {arXiv},
  eprint = {1803.05337},
  url = {https://arxiv.org/abs/1803.05337},
}
```

- Link of the code:- https://github.com/SiddhantGodwani/Computer-Vision-/blob/main/Project_CV_Music_Group25_implementation.pdf

Music Genre Classification Project Documentation

Introduction

This documentation outlines the details of a music genre classification project. The project aims to develop a machine learning model capable of accurately classifying music tracks into different genres based on their audio features. Specifically, this documentation focuses on comparing the performance of two different classification algorithms: K-Nearest Neighbors (KNN) and Naive Bayes.

Purpose

The purpose of this project is to develop a robust machine learning model for music genre classification. By leveraging audio features extracted from music tracks, the model aims to accurately classify each track into one of several predefined genres. This classification capability serves several purposes:

Music Recommendation Systems: Accurate genre classification enables the development of personalized music recommendation systems. By understanding the genre preferences of users, the system can suggest relevant tracks that match their tastes.

Content Organization: Music streaming platforms and digital libraries can benefit from automated genre classification to organize their vast catalogs of music. This helps users discover new music within their preferred genres more efficiently.

Market Segmentation: Understanding the popularity of different music genres among listeners can aid in market segmentation for music industry stakeholders. This information can inform marketing strategies, concert planning, and content creation.

Academic Research: Music genre classification serves as a challenging problem in the field of audio signal processing and machine learning. Research in this area contributes to advancements in both music analysis techniques and classification algorithms.

Objectives

The primary objectives of the project remain unchanged:

Develop a Machine Learning Model: Build a robust machine learning model capable of accurately classifying music tracks into predefined genres.

Utilize Audio Features: Extract relevant audio features from the music tracks to represent their characteristics effectively.

Compare Model Performance: Evaluate and compare the performance of KNN and Naive Bayes classifiers to determine their effectiveness in music genre classification.

Dataset

The project continues to use the "GTZAN Genre Collection" dataset, comprising 1000 audio tracks categorized into 10 different genres.

Feature Extraction and Representation

Feature extraction is a crucial step in music genre classification, as it involves converting raw audio data into a format suitable for machine learning algorithms. The following features are commonly extracted from music tracks:

Mel-Frequency Cepstral Coefficients (MFCCs): Captures the spectral characteristics of the audio signal.

Spectral Centroid: Represents the "center of mass" of the spectrum and provides information about the brightness of the sound.

Chroma Frequencies: Describes the distribution of energy across pitch classes (e.g., C, C#, D, etc.).

Rhythm Patterns: Captures rhythmic information using techniques such as beat tracking and tempo estimation.

These features are then represented as numerical vectors, forming the input data for the classification models.

Model Architecture
K-Nearest Neighbors (KNN) Classifier:
Algorithm: Utilizes a non-parametric approach for classification based on feature similarity.

Training: Stores all training data points in memory.

Inference: Predicts the class label of a test sample by a majority vote of its nearest neighbors in the feature space.

Naive Bayes Classifier:
Algorithm: Based on Bayes’ theorem and assumes independence between features.

Assumption: Assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

Training: Estimates class-conditional probabilities and prior probabilities from the training data.

Inference: Calculates the posterior probability of each class given the input features and selects the class with the highest probability.

Range of Factors in Genres to be Classified
The music genres to be classified encompass a wide range of factors, including but not limited to:

Melody: The sequence of musical notes and their arrangement.

Rhythm: The timing and tempo of musical patterns.

Harmony: The combination of simultaneous musical notes to produce chords and chord progressions.

Instrumentation: The types of musical instruments used in the composition.

Vocal Style: The characteristics of vocal performances, including pitch, timbre, and delivery.

These factors contribute to the distinctiveness of each genre and serve as the basis for classification.

Solution
The proposed solution involves implementing and comparing the performance of KNN and Naive Bayes classifiers for music genre classification:

Feature Extraction: Extract relevant audio features from music tracks, including MFCCs, spectral centroid, and chroma frequencies.

Model Training: Train KNN and Naive Bayes classifiers using the extracted features.

Evaluation: Evaluate the performance of both classifiers using metrics such as accuracy, precision, recall, and F1-score.

Comparison: Compare the performance of KNN and Naive Bayes classifiers to determine their effectiveness in classifying music genres.

By following this approach, the project aims to identify the most suitable classification algorithm for music genre classification, providing insights into the strengths and limitations of each method.

Methodology
The methodology remains consistent with the previous documentation. However, instead of training multiple models with different algorithms, we will focus on implementing and evaluating KNN and Naive Bayes classifiers specifically.

Implementation Details
K-Nearest Neighbors (KNN) Classifier
Implementation: Utilize the KNN algorithm from the scikit-learn library.
Parameters: Tune the number of neighbors (K) to optimize classification performance.
Feature Space: Utilize the extracted audio features as input to the KNN classifier.
Naive Bayes Classifier
Implementation: Implement the Gaussian Naive Bayes algorithm from the scikit-learn library.
Assumption: Assume that features are conditionally independent given the class label.
Feature Space: Utilize the same set of extracted audio features as input to the Naive Bayes classifier.

Evaluation Metrics
The performance of both KNN and Naive Bayes classifiers will be evaluated using the following metrics:

Accuracy: The proportion of correctly classified instances out of the total instances.
Precision: The ratio of correctly predicted positive observations to the total predicted positives.
Recall: The ratio of correctly predicted positive observations to all observations in the actual class.
F1-score: The harmonic mean of precision and recall, providing a balance between the two metrics.

Conclusion
By comparing the performance of KNN and Naive Bayes classifiers on the music genre classification task, we aim to determine the effectiveness of each algorithm in accurately classifying music tracks. This comparison will provide valuable insights into the suitability of different classification approaches for this specific task, ultimately contributing to the successful completion of the project