Human Voice Classification and Clustering

R R Keertana | B.E. CSE

# 1. Introduction

This project focuses on classifying and clustering human voice samples using machine learning techniques. It utilizes extracted audio features to build predictive models and clusters that help identify and analyze human voices. A Streamlit interface is developed for real-time predictions and visual insights.

# 2. Objective and Business Use Cases

• Develop ML models for voice classification and clustering  
• Speaker Identification  
• Gender Classification  
• Speech Analytics for customer service and media  
• Assistive Technologies for accessibility

# 3. Dataset Overview

The dataset consists of various extracted audio features such as spectral centroid, bandwidth, contrast, MFCCs, pitch, energy, and more. Each voice sample is labeled by gender (0 for female, 1 for male).

# 4. Data Preparation

• Missing value handling  
• Normalization using StandardScaler  
• Splitting into train, validation, and test sets

# 5. Exploratory Data Analysis (EDA)

Visualizations used:  
• Pie chart (Gender distribution)  
• Bar chart (Pitch mean per gender)  
• Box plot (MFCC features)  
• Heatmap (Feature correlations)  
• Scatter plot (Pitch vs Energy)

# 6. Clustering

• K-Means and DBSCAN applied  
• Elbow Method to determine optimal clusters  
• Silhouette Score for evaluation  
• Visualization using PCA

# 7. Classification

• Models used: Random Forest, SVM, Neural Networks  
• Hyperparameter tuning and feature selection  
• Evaluation using Accuracy, Precision, Recall, F1-score, and Confusion Matrix

# 8. Streamlit Application

• Introduction page with project image  
• EDA page with 5 questions and chart types  
• Prediction form for 10 input features with 9-digit decimal range  
• Clustering page with elbow plot, score, and visualization  
• About Me page with developer introduction

# 9. Future Scope

• Integrate speaker identification using raw audio  
• Apply deep learning techniques  
• Enable real-time voice input in the app  
• Expand to multilingual voice analysis

# 10. Conclusion

This project demonstrates the power of machine learning in understanding and analyzing human voice patterns. From clustering to real-time prediction using Streamlit, the complete pipeline showcases voice-based AI applications.