# Project Report: Human Voice Classification and Clustering

## Project Title

Human Voice Classification and Clustering

## Skills Gained

- Python Programming

- Data Preprocessing

- Feature Engineering and Selection

- Clustering (KMeans, DBSCAN)

- Classification (Random Forest, SVM, MLP)

- Model Evaluation (Silhouette Score, Classification Report)

- Audio Processing with Librosa

- Streamlit Web App Development

## Domain

Speech Processing | Machine Learning | Web Application

## Problem Statement

Develop a machine learning-based system to classify and cluster human voice samples based on extracted audio features. The system should:

- Preprocess a dataset of voice recordings,

- Apply clustering algorithms to discover natural groupings,

- Train classification models to predict gender from voice features,

- Enable real-time prediction on uploaded audio using a web interface.

## Dataset Used

vocal\_gender\_features\_new.csv

- Pre-extracted statistical audio features

- Features: Mean Frequency, Spectral Entropy, Skewness, Kurtosis, Modulation Index, etc.

- Target label: 0 = Female, 1 = Male

## Technologies Used

- Languages: Python

- Libraries: pandas, numpy, scikit-learn, librosa, matplotlib, seaborn, joblib

- Tools: Streamlit, Librosa, Joblib

## Exploratory Data Analysis (EDA)

- Summary Statistics

- Gender Distribution

- Correlation Heatmap

- Boxplots & KDE Plots

## Clustering Techniques

KMeans:

- Clustered data using KMeans with user-defined K

- Silhouette Score for evaluation

- PCA for visualization

DBSCAN:

- Density-based clustering

- Parameters: eps, min\_samples

- Identified noise and arbitrary clusters

## Classification Models

Models:

1. Random Forest

2. Support Vector Machine (SVM)

3. Multilayer Perceptron (MLP)

Metrics:

- Accuracy

- Precision

- Recall

- F1-score

## Live Voice Prediction

- Upload WAV file

- Extract audio features via Librosa

- Classify as Male or Female using trained models

- Display waveform and prediction

## Web Application with Streamlit

Modules:

1. Home

2. EDA

3. Clustering

4. Classification

5. Audio Prediction

6. Conclusion

## Key Results

- Random Forest achieved ~95% accuracy

- KMeans provided clear visual clusters with PCA

- Real-time audio classification worked effectively

## Conclusion

- Statistical features effectively used for clustering and classification

- Streamlit provided interactive interface

- Models achieved high accuracy

## Future Enhancements

- Use deep learning (CNNs)

- Expand dataset

- Add multi-class classification

- Deploy with API access

## Appendix

Model Files:

- random\_forest\_model.pkl

- svm\_model.pkl

- mlp\_model.pkl

- scaler.pkl

Files:

- app.py

- vocal\_gender\_features\_new.csv