SPEACH RECOGNITION SYSTEM USING MACHINE LEARNING

Presented by:

1. **Khondker Sayif Ali(2111323642)**
2. **Md. Injamam ul Haque(2111598642)**
3. **Nashita Tasneem Noor(2132126642)**

INTRODUCTION

Aim: Build a speech recognition system using ML

Input: Human voice dataset (10 speakers)

Output: Predict correct speaker

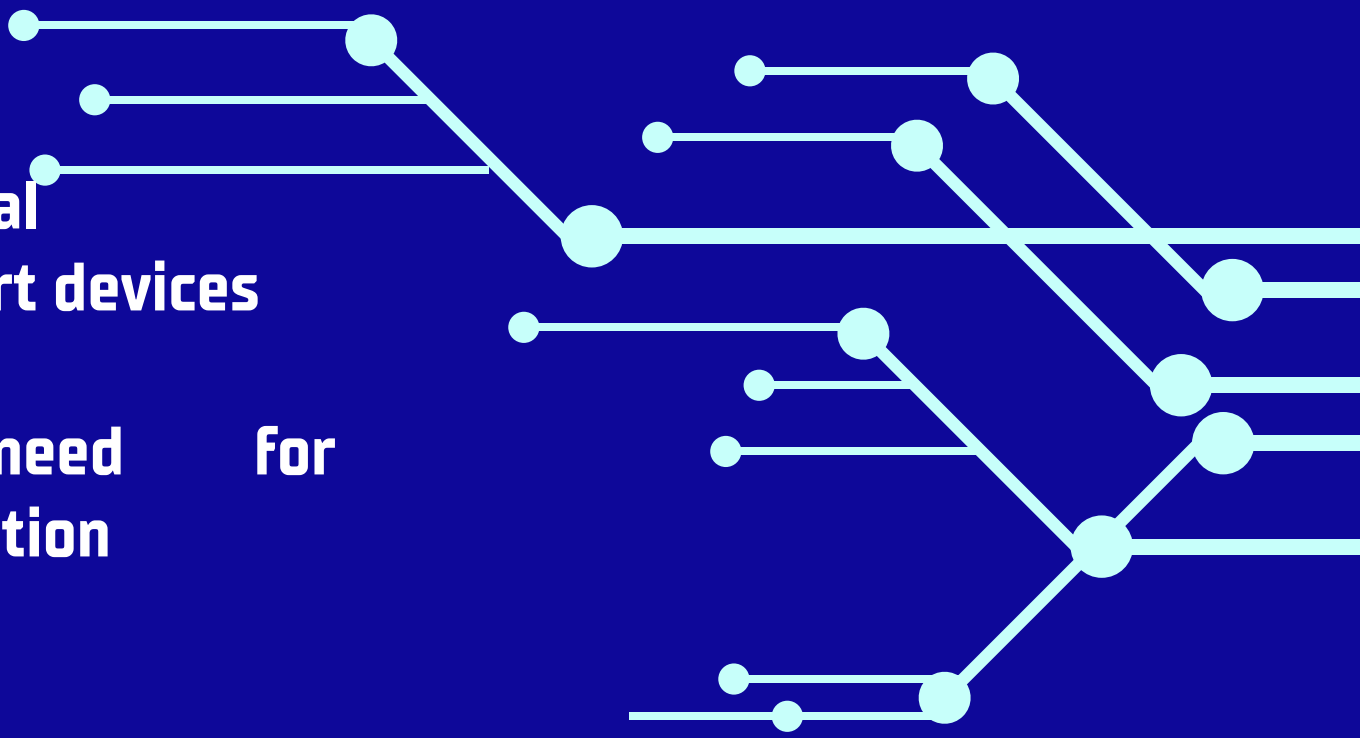
Tools: Python, Librosa, Scikit-learn, Google Colab

MOTIVATION

Voice-based authentication is growing

·Applications: Security,personal assistants, smart devices

Real-world need for speaker recognition

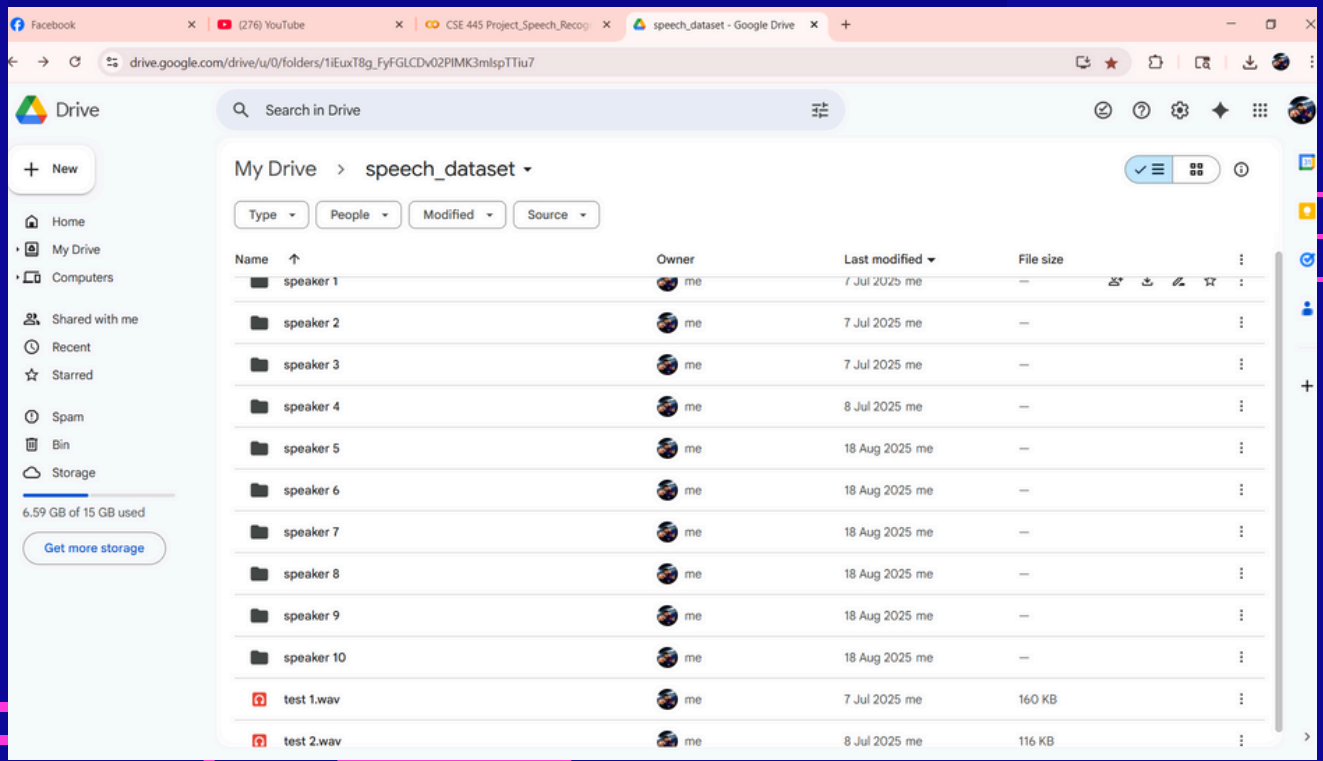


DATASET OVERVIEW

Dataset contains 10 different speakers

Each speaker contributed multiple audio samples· Format: WAV files

Features extracted using MFCCs



PREPROCESSING

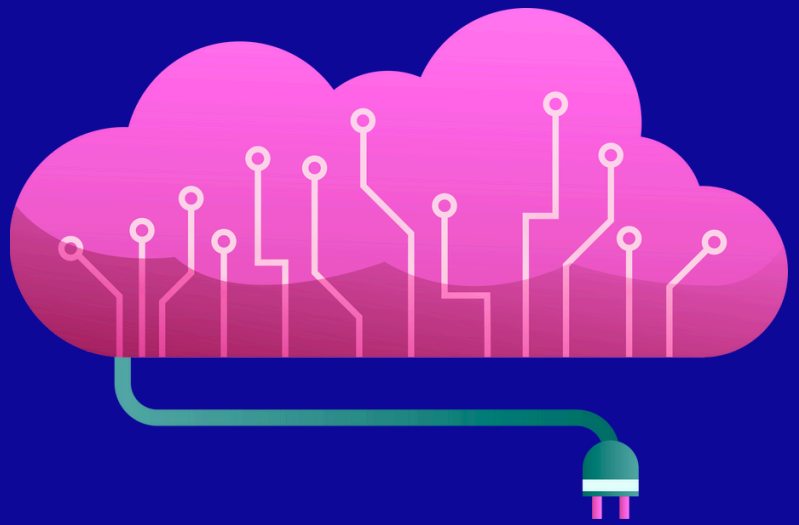
1. Audio data cleaned and normalized
2. Converted to MFCC feature vectors
3. Labels encoded using LabelEncoder
4. Train-Test split applied

FEATURE EXTRACTION

- MFCC (Mel-Frequency Cepstral Coefficients)
- Captures timbral texture of audio
- Widely used in speech recognition

MODEL USED

- Model – Logistic Regression
- Simple yet effective baseline model
- Converts MFCC features to numerical vectors
- Learns class boundaries between speakers

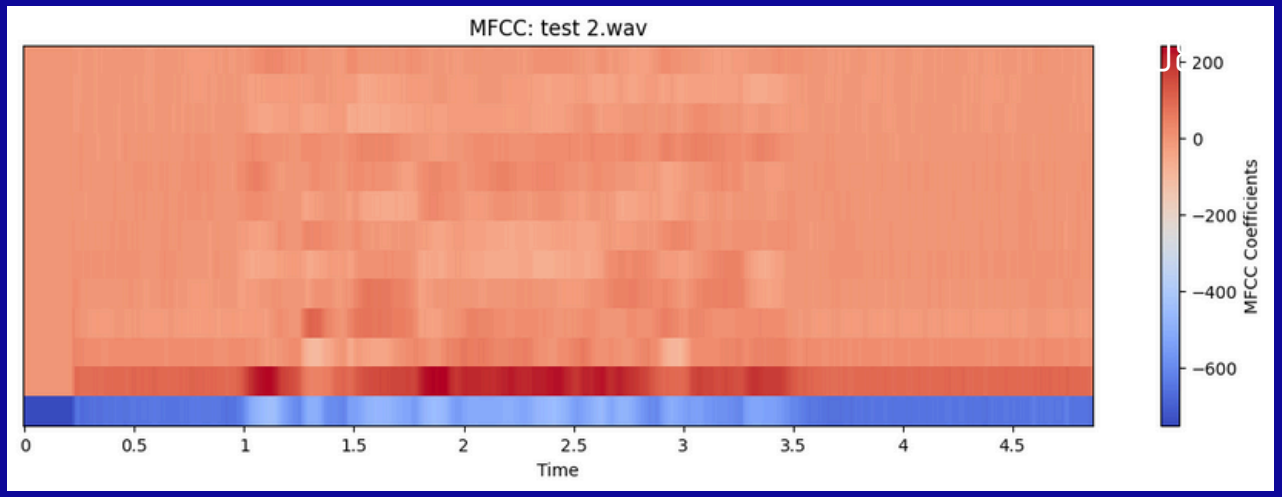


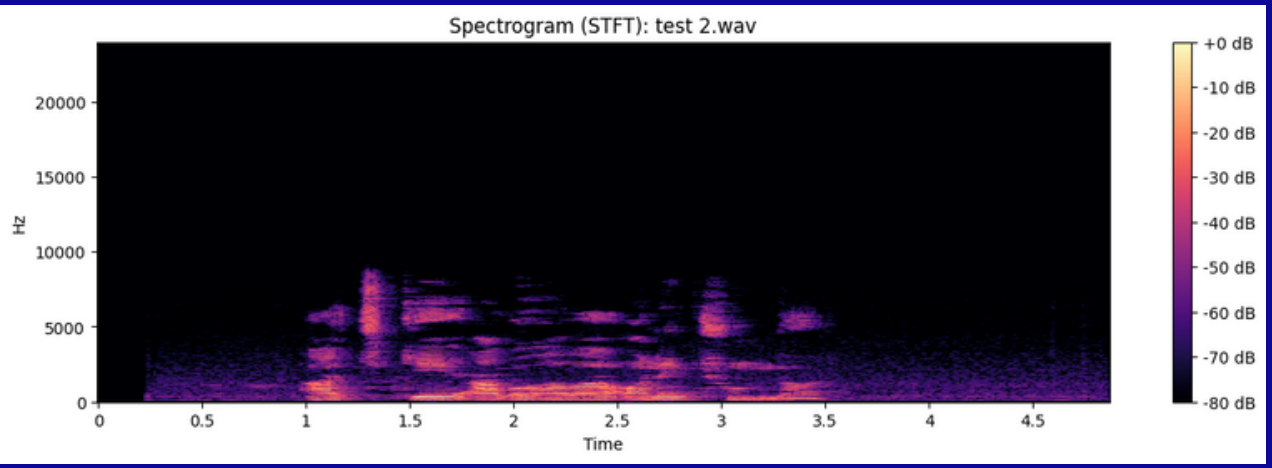
CODE PIPELINE

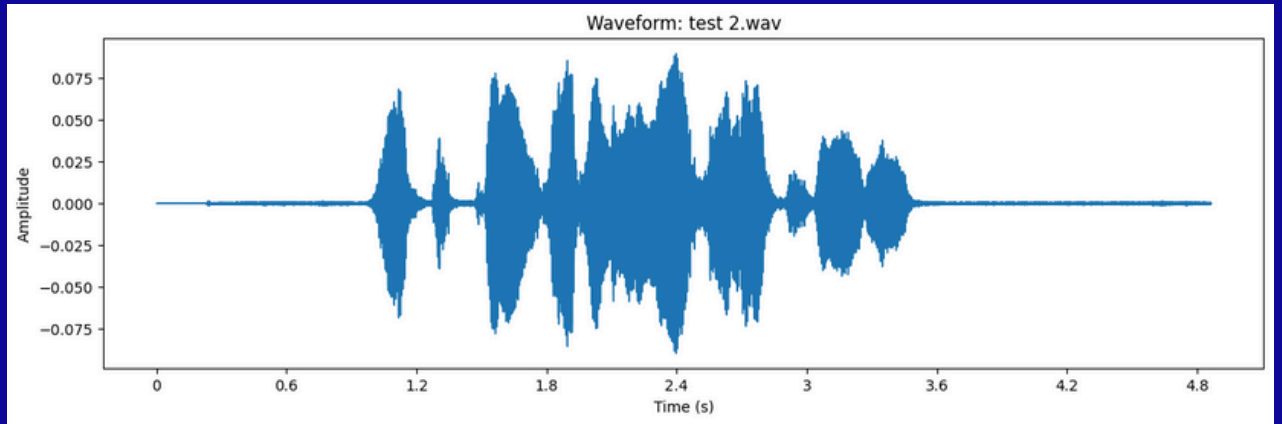
Sequentiially

- LOAD DATASET
- PREPROCESS AUDIO
- EXTRACT MFCC FEATURES
- TRAIN ML MODEL
- EVALUATE WITH ACCURACY & CLASSIFICATION REPORT

MFCC SPECTROGRAM







EVALUATION METRIC

Classification Report used for detailed analysis

Accuracy: 50%

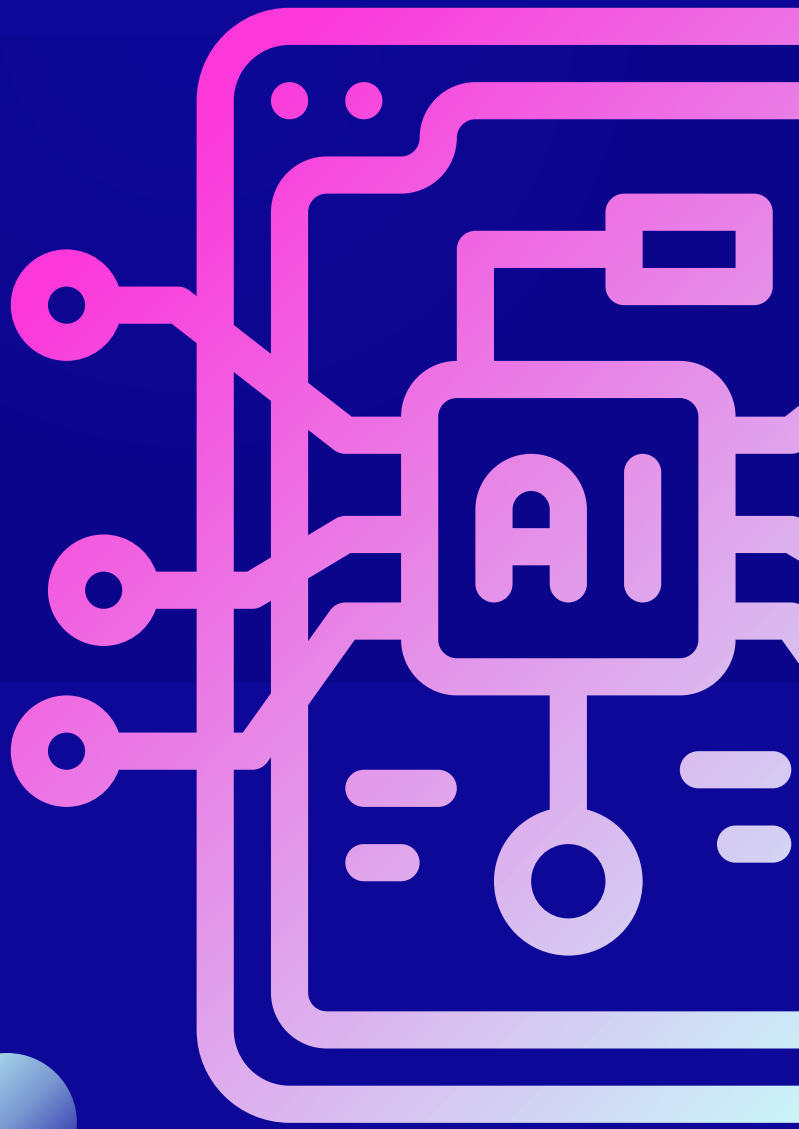
Metrics: Precision, Recall, F1-score

CLASSIFICATION REPORT

Example Results:

- **Precision (avg): 0.52**
- **Recall (avg): 0.50**
- **F1-score (avg): 0.49**

Shows balanced but limited performance



OBSERVATIONS & CHALLENGES

- Limited dataset size (10 speakers)
- Accuracy drops with similar voices
- Background noise affected results
- Logistic Regression not robust

FUTURE WORK

Increase dataset size (more speakers)

- Use deep learning models (CNN, RNN)
- Apply noise reduction techniques
- Deploy as real-time voice recognition app

CONCLUSION

Built a simple speaker identification system

Preprocessed voices & extracted MFCCS

Trained Logistic Regression model

Results show promise but need improvement

THANK YOU

Questions?

Presented by:

1. **Khondker Sayif Ali(2111323642)**
2. **Md. Injamam ul Haque(2111598642)**
3. **Nashita Tasneem Noor(2132126642)**