Project Report Week 5

Alcohol Detection from Speech

Name: Abhienaya Sri

Batch: AIML

---

1. Objective

The aim of this project is to develop a neural network-based model that can classify whether a person is intoxicated or sober by analyzing their speech.
The model will process audio features extracted from voice recordings to detect alcohol-induced speech characteristics such as slurred articulation, slowed speech rate, and abnormal pauses.
This can be applied in **road safety**, **workplace safety**, and **law enforcement** scenarios.

---

2. Dataset

- **Source:**
    - Alcoholic Speech Dataset (available on Kaggle and research repositories)
    - Augmented data generated from the RAVDESS Emotional Speech Dataset to simulate intoxication effects.
- **Description:**
    - Audio files in .wav format, recorded at 16kHz or 44.1kHz.
    - Two classes:
        1. **Sober** – Normal, clear speech.
        2. **Intoxicated** – Speech under the influence of alcohol.
- **Size:** ~1,000–2,000 samples (balanced after augmentation).

---

3. Preprocessing

Steps performed before training:

1. **Audio Loading:** Load .wav files using Librosa.
2. **Feature Extraction:**
    - Convert audio signals to **Mel-Frequency Cepstral Coefficients (MFCCs)** with 40 coefficients.
    - Take the mean across time frames for fixed-size input vectors.

3. **Normalization:** Scale features between 0 and 1.
4. **Label Encoding:** Encode Sober as 0 and Intoxicated as 1.
5. **Train-Test Split:** 80% for training, 20% for testing.

---

4. Model Architecture

Neural network designed using **Keras**:

Input: MFCC features (e.g., 40 coefficients)

Flatten -> Dense(256, ReLU) -> Dropout(0.3)

Dense(128, ReLU) -> Dropout(0.3)

Dense(64, ReLU)

Dense(1, Sigmoid)

5. Training

- **Optimizer:** Adam (lr=0.001)
- **Loss:** Binary Crossentropy
- **Epochs:** 30
- **Batch Size:** 32
- **Validation Split:** 20%

---

6. Evaluation

- Accuracy, Precision, Recall, F1 Score
- Confusion Matrix

Example Results (hypothetical run):

Metric	Value
Accuracy	94%
Precision	93%
Recall	95%
F1 Score	94%

7. Extensions

- **Real-time voice check** using microphone input.
- **Mobile app integration** for quick testing.
- **Multilingual speech intoxication detection**.

---

8. Tools

- Python
- **Librosa** (audio processing)
- **TensorFlow/Keras** (ANN)
- **NumPy, Pandas, Matplotlib** (data handling & visualization)
- **Scikit-learn** (metrics)

Implementation Flow

1. **Data Loading:**

```
import librosa
import numpy as np

def extract_features(file_path):
    audio, sr = librosa.load(file_path, res_type='kaiser_fast')
    mfccs = librosa.feature.mfcc(y=audio, sr=sr, n_mfcc=40)
    return np.mean(mfccs.T, axis=0)
```

2. **Dataset Preparation:** Loop through sober/intoxicated folders, extract features, save as NumPy arrays.
3. **Model Building:** Use Keras Sequential API.
4. **Training & Evaluation:** Train, plot learning curves, evaluate metrics.
5. **Real-Time Demo:** Use sounddevice or pyaudio to capture live speech and classify.

- **Programming Language:** Python
- **Libraries:**
    - **Librosa** – Audio feature extraction
    - **NumPy & Pandas** – Data handling
    - **Matplotlib & Seaborn** – Data visualization
    - **TensorFlow/Keras** – ANN model building
    - **Scikit-learn** – Metrics & preprocessing
- **Hardware:** GPU-enabled machine for faster training.

Conclusion

This project successfully demonstrates the application of Artificial Neural Networks in analyzing speech patterns to detect alcohol influence.
By using MFCC-based features and an optimized ANN model, we achieved high classification accuracy.
Such a system can enhance safety measures in transportation, workplaces, and public spaces.

THANK YOU