MemoTag AI/ML Cognitive Voice Task – April 2025

Author: Izhan Abdullah
Date: April 2025

Project Title: Voice-Based-Cognitive-Decline-Pattern-Detection

---

1. Objective

To build a proof-of-concept system using audio speech recordings to detect patterns indicative of cognitive stress or decline. The goal is to extract both acoustic and linguistic features from voice samples and apply unsupervised machine learning techniques to identify anomalies that could serve as early warning signs of cognitive impairment.

---

2. Problem Statement

As part of MemoTag's speech intelligence module, the task involves processing 5-10 anonymized voice clips. The aim is to extract both speech-based and linguistic features, such as hesitation markers, pitch variability, word recall, and speech rate, to detect patterns that may suggest early cognitive decline. This project focuses on identifying cognitive impairment indicators using unsupervised learning approaches.

---

3. Methodology

3.1 Audio Preprocessing

- **Format**: Audio files in .wav format.
- **Tools**:
    - **Librosa** for audio processing and feature extraction.
    - **SpeechRecognition** for transcription of the audio into text.

3.2 Feature Extraction

- **Acoustic Features**:
    - **Duration**: Total length of the voice sample.
    - **Avg_pitch**: Mean voice pitch to assess overall voice tone.
    - **Pitch_var**: Pitch variability, a potential stress indicator.
    - **Pause_count**: Number of silent segments in speech.
- **Linguistic Features**:
    - **Word_count**: Total number of words spoken.

- **Hesitations**: Frequency of hesitation markers ("uh", "um", etc.).
- **Avg_word_length**: Linguistic complexity (average length of words).

3.3 Unsupervised Learning Approach

- **Model**: **KMeans Clustering** to detect abnormal patterns in speech behavior.
- **Dimensionality Reduction**: Principal Component Analysis (PCA) for 2D visualization of features.
- **Goal**: To detect abnormal speech patterns and group samples into clusters (normal vs. unusual speech).

---

4. Visualizations

- **Clustering Visualization**:
  A PCA scatter plot was created to show how the audio samples cluster based on the extracted features. The plot clearly differentiates between normal and abnormal speech patterns, highlighting potential signs of cognitive decline.

---

5. Key Insights

- **Hesitation Frequency**: Hesitation markers such as "uh" and "um" were the most prominent indicators of cognitive stress. Higher frequency of these markers suggests difficulty in word recall.
- **Pitch Variability**: Variability in pitch was also a significant indicator. Increased pitch variation may suggest cognitive strain or anxiety.
- **Pause Count per Sentence**: A higher number of pauses per sentence was indicative of mental hesitation, a potential sign of cognitive decline.
- **Word Complexity**: Linguistic complexity, measured by the average word length and complexity, provided secondary insights, with simpler vocabulary linked to potential cognitive impairment.

---

6. Model Evaluation

- **KMeans Clustering**:
  The unsupervised clustering method was chosen due to its simplicity and interpretability. It effectively grouped speech samples without requiring labeled data, which is typically unavailable in clinical settings.
  - **Advantages**:
    - Lightweight and interpretable.

        - Ideal for early-stage exploration of cognitive decline.
        - Does not require labeled data for training, which is advantageous given the limited availability of labeled cognitive impairment data.
    - **Limitations**:
        - The model may not fully capture complex cognitive patterns without further data or more advanced techniques like deep learning.
        - Results depend heavily on the quality of the feature engineering process.

---

7. Next Steps

To enhance the clinical relevance of this system:

- **Data**: Collect more labeled data from real patients to train a more robust model.
- **Task-based Assessments**: Incorporate memory and naming tasks to better assess cognitive function.
- **Deep Learning**: Use advanced NLP models (e.g., **BERT**) to capture more nuanced speech patterns, such as sentence gaps and semantic meaning.
- **Collaboration with Neurologists**: Work closely with medical experts to refine the feature selection and model development process for greater clinical applicability.

---

8. Deliverables

- **Python Notebook (.ipynb)**: Contains the full code for audio processing, feature extraction, and clustering.
- **Feature CSV**: A CSV file containing the extracted features from the audio samples.
- **Clustering Plot**: A PCA scatter plot visualizing the clustering of speech samples.
- **Report (Markdown & PDF)**: The final written report explaining the methodology, insights, and next steps.

---

9. Conclusion

This proof-of-concept system demonstrates the potential of using unsupervised learning techniques and voice analysis to detect cognitive stress and impairment. The results, while promising, require further refinement with real-world data and collaboration with healthcare professionals to move towards clinical application.

Author: [Izhan Abdullah]
MemoTag AI/ML Cognitive Voice Task
April 2025