ML MODEL FOR EARLY COGNITIVE IMPAIRMENT DETECTION

-Aarushi Anand

This analysis combines person-specific rule-based thresholds with SVM classification to detect cognitive impairment risk. Below are the key findings and methodological insights:

1. Most Insightful Features

The notebook extracts several speech-based features that are relevant for identifying potential cognitive impairment indicators. The most insightful among these include:

- **Pauses per minute**: Long pauses (>1.2s) were counted and normalized over time, as increased frequency of such pauses is often linked to cognitive difficulties.

- **Repetition rate**: Both individual word and short phrase repetitions were flagged. Higher repetition frequency may reflect memory or speech formulation issues.

- **Speech rate (WPM)**: Deviations from a typical word-per-minute range may suggest cognitive or motor issues.

- **Pitch variability (standard deviation)**: Monotonic speech (low pitch variance) may be indicative of affective or neurological disorders.

- **Incomplete sentence rate**: Reflects syntactic coherence and potential difficulties in expression.

These features are directly interpretable and align well with known clinical markers in neuropsychological assessments.

2. ML Methods Used and Rationale Rule-Based Risk Classification:

- The initial approach uses an if-else scoring system to classify individuals into Low, Moderate, or High Risk. This rule-based system evaluates the extracted features against predefined thresholds. It offers transparency and interpretability, which are beneficial in a clinical context.

- SVM Classifier for Generalization: In the later part of the notebook (not shown above but inferred), a Support Vector Machine (SVM) is employed to generalize categorization across multiple speakers. SVMs are robust with small datasets and effective in high-dimensional spaces, making them suitable for this scenario where feature vectors are compact but potentially non-linearly separable.

3. Sample visualizations of feature trends

1. Box Plot has been used to visualize the distribution of a key feature (e.g., mfcc_0, speech_rate) for each class (Normal vs. Risk)
2. Correlation Heatmap: Shows how features relate to each other and to the label.

4. Potential Next Steps Toward Clinical Robustness

Hesitation rate: Hesitative utterances like “uh”, “um”, “hmm” were tracked as proxies for verbal disfluency.

Several models were tried like Nemo(by Nvidia), whisper, speech matic, Krisp etc but due to inefficiencies in the transcribed text they couldn't be used so the approach was finally boiled down to using Assembly AI which gave the best results.

Other tools and models can be researched upon and fine-tuned for the same cause and tried in future.

To evolve this prototype into a clinically reliable tool, the following steps are recommended:

a. Data Scaling

- **Increase dataset size**: More speakers, age diversity, and validated clinical labels (diagnosed vs. control) are essential.

- **Balance demographics**: Gender, accent, age, and language proficiency need to be controlled or accounted for.

b. Model Enhancement

- **Feature engineering**: Include linguistic complexity (e.g., syntactic depth, lexical diversity) and acoustic biomarkers (e.g., jitter, shimmer).
- **Multimodal modeling**: Combine audio features with transcript-based NLP features for a richer profile.
- **Deep learning models**: Test CNNs/RNNs on spectrograms or use transformer-based embeddings (e.g., wav2vec) for improved representation.

c. Validation Strategy

- **Cross-validation**: Employ stratified k-fold validation to ensure model generalizability.
- **Clinical benchmarking**: Compare against neuropsychological assessments (e.g., MMSE, MoCA).
- **Human-in-the-loop testing**: Involve clinicians for interpretability validation and iterative improvement.

d. Deployment Considerations

- **Privacy and security**: Ensure compliance with HIPAA/GDPR for audio data.
- **Real-time processing**: Optimize for latency and streaming inputs.
- **Explainability**: Use SHAP or LIME to explain model predictions to clinicians.