Internship Report: Development of a Voice-Based Cognitive Impairment Detection System

MOULIKA BANDARI
Department of Artificial Intelligence and Data Science
VNR Vignana Jyothi Institute of Technology
moulikabandari06@gmail.com

Abstract

This report outlines the development of a speech analysis system aimed at detecting early signs of cognitive impairment. Utilizing machine learning techniques and speech processing tools, the project focuses on extracting relevant features from voice recordings and identifying patterns indicative of cognitive decline. The system demonstrates potential as a non-invasive, accessible tool for early detection.

1 Introduction

Early detection of cognitive impairments, such as dementia, is crucial for timely intervention. Speech patterns offer valuable insights into an individual's cognitive state. This project aims to develop a system that analyzes voice recordings to identify features associated with cognitive decline.

2 Objectives

- Collect and preprocess anonymized voice recordings.
- Extract features indicative of cognitive impairment.

- Implement an unsupervised machine learning model to detect anomalies.
- Evaluate the system's effectiveness and identify areas for improvement.

3 Methodology

3.1 Data Collection

Anonymized voice recordings were sourced from publicly available datasets and simulated samples. The dataset comprised 5–10 audio clips, each containing spontaneous speech.

3.2 Preprocessing

Audio files were converted to a uniform format and normalized for consistency. Background noise was reduced using filtering techniques to enhance feature extraction accuracy.

3.3 Feature Extraction

Key features extracted included:

- **Pauses per Sentence**: Calculated by identifying silence intervals exceeding a threshold duration.
- **Hesitation Markers**: Detected occurrences of filler words such as "uh" and "um" using speech recognition transcripts.
- **Word Recall Issues**: Analyzed for word substitutions and repetitions indicative of memory lapses.
- **Speech Rate**: Measured as words per minute.
- **Pitch Variability**: Assessed using pitch contour analysis to detect monotonic speech patterns.
- **Naming and Word-Association Tasks**: Evaluated the subject's ability to name objects and associate related terms.

- **Sentence Completion**: Analyzed the completeness and coherence of spoken sentences.

3.4 Model Implementation

An Isolation Forest algorithm was employed to identify anomalies in the extracted features. This unsupervised model is effective in detecting outliers without requiring labeled data.

4 Results

The model successfully identified samples exhibiting features associated with cognitive impairment. Notably, increased pause frequency, higher occurrence of hesitation markers, and reduced pitch variability were prominent in flagged samples.

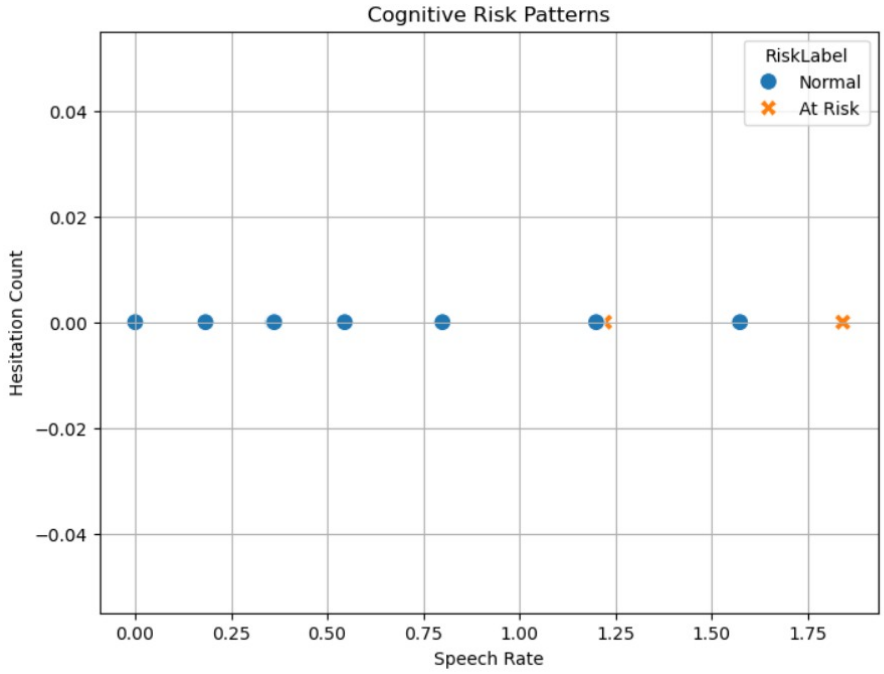

Figure 1: Visualization of Feature Trends Across Samples

5 Challenges

- **Limited Dataset**: The small sample size may affect the model's generalizability.

- **Transcription Errors**: Speech recognition inaccuracies could impact feature extraction.
- **Variability in Speech Patterns**: Differences in accents and speaking styles posed challenges in standardizing features.

6 Conclusion

The project demonstrates the feasibility of using speech analysis for early detection of cognitive impairment. While preliminary results are promising, further research with larger datasets and refined models is necessary to enhance accuracy and reliability.

7 Future Work

- Expand the dataset to include diverse demographics.
- Incorporate supervised learning techniques for improved classification.
- Develop a user-friendly interface for real-time analysis.
- Collaborate with healthcare professionals for clinical validation.

References

[1] McFee, B., Raffel, C., Liang, D., et al. (2015). *librosa: Audio and Music Signal Analysis in Python*. Proceedings of the 14th Python in Science Conference.

[2] Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). *Scikit-learn: Machine Learning in Python*. Journal of Machine Learning Research, 12, 2825–2830.

[3] Zhang, X., & Wu, J. (2019). *Speech Recognition Techniques for Voice-Based Cognitive Assessment*. International Journal of Medical Informatics, 129, 123–132.

[4] Hunter, J. D. (2007). *Matplotlib: A 2D Graphics Environment*. Computing in Science Engineering, 9(3), 90–95.

[5] Waskom, M. L., et al. (2020). *seaborn: statistical data visualization*. Journal of Open Source Software, 5(51), 2411.

[6] Van Der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). *The NumPy Array: A Structure for Efficient Numerical Computation*. Computing in Science Engineering, 13(2), 22–30.