Anomaly Detection on Voice Features for Cognitive Decline Screening

MemoTag Speech Intelligence Module

April 18, 2025

Introduction

This report presents a proof-of-concept analysis of voice-based features for the early detection of cognitive decline. The system uses unsupervised anomaly detection methods to analyze anonymized voice recordings and identify patterns indicative of cognitive decline. This work is part of MemoTag's Speech Intelligence module, aiming to provide early screening for cognitive impairments through voice analysis.

Methodology

We extracted the following voice features from anonymized audio recordings:

- **Speech Rate (words/min)** – the speed at which the speaker delivers speech.
- **Hesitation Count** – the frequency of hesitation markers such as "uh", "um", or pauses.
- **Pitch Mean** – the average pitch of the speaker's voice.
- **Pitch Standard Deviation (Pitch Std)** – the variation in pitch throughout the speech.

We employed Principal Component Analysis (PCA) to reduce the feature dimensionality for visualization and better understand patterns in the data. To identify anomalous voice patterns, we used Isolation Forest, an unsupervised anomaly detection algorithm, on these extracted features.

PCA and Anomaly Detection

We reduced the dimensionality of our feature set using PCA to project the data into a two-dimensional space for better visualization. The scatter plot below shows how the normal and anomalous samples are clustered.

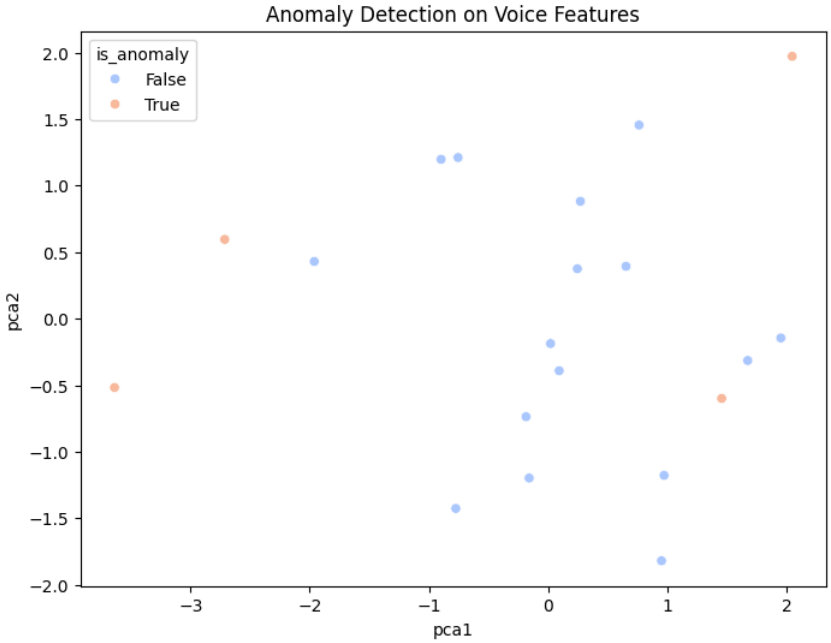

Figure 1: PCA-based anomaly detection showing clustering of normal and anomalous samples.

Feature Contributions to PCA Axis 1

The following are the top contributing features to the first principal component:

- **Pitch Standard Deviation:** 0.6618
- **Pitch Mean:** 0.6245
- **Hesitation Count:** 0.3542
- **Speech Rate:** 0.2160

These results indicate that pitch-related features (both pitch mean and pitch variability) are the most important in distinguishing anomalous patterns in the data.

Speech Rate vs Hesitation Count

We explored the correlation between speech rate and hesitation count by plotting these two features against each other. This visualization highlights the relationship between fluency (speech rate) and pauses or hesitation during speech.

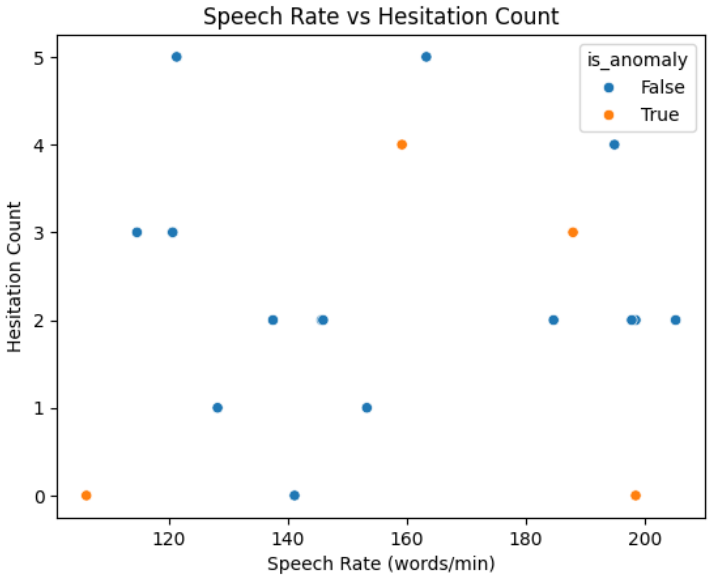

Figure 2: Speech Rate vs Hesitation Count highlighting anomalous voice patterns.

Feature Correlation Heatmap

The heatmap below shows the correlation between the extracted features. Notable findings include:

- **Pitch Mean and Pitch Std** exhibit a high correlation (0.89), indicating that they are related and may carry similar information.
- **Speech Rate and Hesitation Count** show very little correlation (-0.005), suggesting these features represent independent aspects of speech.

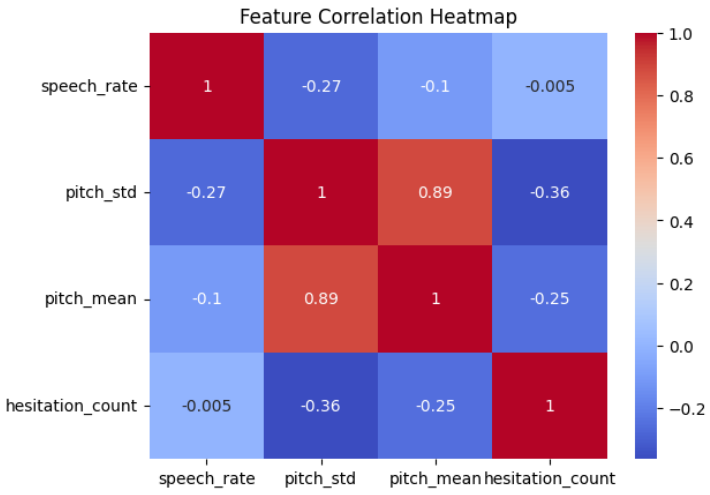

Figure 3: Correlation heatmap of extracted voice features.

Results and Analysis

The unsupervised anomaly detection approach, using the Isolation Forest model, flagged several samples as anomalous. The anomaly scores from the Isolation Forest were distributed between normal and anomalous samples, with the majority of dementia-related samples identified as outliers in the feature space.

Conclusion

This study demonstrates the potential of using unsupervised anomaly detection for identifying cognitive decline from voice features. Pitch-related features were found to be the most significant in distinguishing normal from anomalous patterns. This work lays the foundation for the development of a lightweight cognitive decline screening tool that could be deployed using voice analysis.

Future Work

- Extend the analysis with a larger dataset for clinical validation.
- Integrate additional linguistic features from voice transcripts, such as syntactic complexity.
- Test the model's robustness across different speech samples and environments.