Chapter 5: Results

5.1 Classification Performance

Overall Results

Task Algorithm Accuracy F1 Score

Key Finding: Interview Outperforms Reading

The most striking result is the ~15 percentage point performance gap between tasks:

This suggests spontaneous speech contains richer depression markers than controlled

Classifier Comparison

Random Forest marginally outperformed SVM:

Larger advantage on interview data may reflect RF's ability to capture complex feature

5.2 Feature Importance Analysis

Reading Task Top 10 Features

Rank Feature Importance

Interpretation - Reading Task:

- **Spectral Slope Features:** Top feature (slopeUV0-500) measures spectral tilt in unvoiced regions.
- **Temporal Features:** Loudness peaks/sec and voiced segment variability indicate rhythm and prosodi

Interview Task Top 10 Features

Rank Feature Importance

Interpretation - Interview Task:

- **Spectral Flux Variability:** Most important feature! Measures frame-to-frame spectral change varia
- **Voice Quality Measures:** Hammarberg index and alpha ratio reflect breathiness and vocal strain. T
- **Pausing Behaviour:** Mean and stddev of unvoiced segment length (pauses/hesitations). Clinically m
- **Loudness Dynamics:** Variability of loudness slopes indicates prosodic expression. Reduced modulat

5.3 Task Comparison

Feature Overlap

Comparing top 10 features reveals limited overlap:

- **Shared:** Only mfcc1V (mean and variability)

Different acoustic markers are salient depending on speech context.

Dominant Feature Categories by Task

Category Reading Interview

Clinical Interpretation

Reading Task: Reveals voice production characteristics (spectral slope, MFCC) and basic rhythm.

Interview Task: Reveals cognitive and affective processes:

Interview places greater demands on executive function, emotional regulation, language

5.4 Visualisations

Feature importance plots reveal:

- Reading task: importance more evenly distributed across features

See: figures/reading_feature_importance.png, figures/interview_feature_importance.png

5.5 Statistical Significance

Cross-validation standard deviations (~5-7%) indicate reasonably stable results. The 15-point

5.6 Summary of Findings

1. Interview speech is more informative: 87% vs 72% accuracy

These findings directly address the research question with interpretable insights into how

Estimated length: 6-8 pages