

Chapter 1: Introduction

Restructured to match A-grade dissertation format

1.1 Motivation

Depression affects **over 280 million people globally**, making it one of the leading causes of disability worldwide [WHO, 2024]. In the United Kingdom alone, approximately **one in six adults** experiences depression or anxiety in any given week, with an estimated economic burden of **£105 billion annually**.

Despite its prevalence, depression remains significantly **underdiagnosed and undertreated**. Current diagnostic methods rely primarily on clinical interviews and self-report questionnaires such as the Patient Health Questionnaire (PHQ-9). These approaches face several critical limitations:

- **Subjectivity:** Diagnostic agreement between clinicians can be inconsistent, with inter-rater reliability as low as 0.28 kappa for certain depression subtypes.
- **Self-report bias:** Patients may underreport symptoms due to stigma, lack of insight, or difficulty articulating their experiences.
- **Resource constraints:** Clinical assessments require trained mental health professionals, limiting scalability and accessibility.
- **Access barriers:** Many individuals, particularly in underserved communities, lack access to specialist care.

These limitations have motivated research into **objective, scalable biomarkers** for depression. Among the most promising candidates is **speech**. Depression affects cognitive processes that directly influence speech production—changes that are often involuntary and difficult to consciously mask. Recording and analysing speech is non-invasive, inexpensive, and can be performed remotely, making it an attractive option for large-scale screening.

Recent advances in machine learning have enabled automatic detection of depression from speech with accuracy comparable to general practitioners (~80%). However, most research has focused on maximising classification accuracy, treating models as "black boxes." This raises a critical question:

Which aspects of speech actually indicate depression?

Understanding which acoustic features predict depression is essential for:

1. **Clinical utility:** Clinicians cannot act on "83% accuracy"—they need interpretable markers.
2. **Scientific understanding:** Identifying robust biomarkers advances our understanding of how depression manifests in behaviour.
3. **System design:** Knowing which features matter informs what data to collect and how to deploy screening tools.

This dissertation addresses this gap by systematically analysing which acoustic features are most predictive of depression, and how their importance differs between controlled (read) and naturalistic (spontaneous) speech.

1.2 Aims

This project aims to:

1. **Identify predictive features:** Determine which acoustic and prosodic features are most strongly associated with depression, using interpretable machine learning techniques.
2. **Compare speech tasks:** Analyse whether the same features predict depression in read speech (controlled phonetic content) versus spontaneous speech (natural conversation), and which task yields more reliable detection.
3. **Provide interpretable analysis:** Move beyond accuracy metrics to explain *why* models predict depression, using feature importance methods such as SHAP values.

Research Question

"Which acoustic features of speech are most predictive of depression, and how do they differ between read and spontaneous speech?"

1.3 Outline

This dissertation is structured as follows:

Chapter	Title	Content
2	Background	Depression & speech, acoustic features, ML approaches, related work
3	Design	Methodology, data selection, experimental approach
4	Implementation	Technical details, tools, pipeline
5	Evaluation	Results, feature rankings, task comparison
6	Discussion	Critical analysis, limitations, implications
7	Conclusion	Summary and future work

This follows the structure of FATA, IDA, and GIST (A-grade dissertations)

Introduction = WHY (motivation, aims, outline)

Background = WHAT (existing work, theory, technology)