Digital Systems and Design

Final Project Plan

Signal Processing and Machine Learning for Wrist Pulse Waveforms in Arterial Stiffness Estimation

Dinh Bach Tue & 101733959
Supervisor: PhD Researcher Bahmani Nima

Introduction & Design Problem

Cardiovascular diseases are a leading cause of death worldwide, and arterial stiffness—measured as central pulse wave velocity (PWV)—is a well-established marker for cardiovascular risk. However, clinical measurement of arterial stiffness requires specialized, non-portable equipment and is not widely accessible.

Motivation:
With the rise of wearable devices capable of measuring wrist pulse waveforms (PPG/pressure), there is an opportunity to enable non-invasive, accessible, and continuous vascular health monitoring outside clinical settings. Despite this potential, the precise relationship between central arterial stiffness and the morphology of wrist pulse waveforms is not fully understood, especially in real-world (noisy, wearable) contexts.

Design Problem:
This project aims to bridge that gap by combining advanced signal processing and machine learning techniques to (1) understand how arterial stiffness is reflected in wrist pulse waveforms using a large synthetic (simulated) database, and (2) apply the developed methods to real wrist PPG data from wearable devices for practical arterial stiffness estimation. The work will address how to robustly extract and interpret stiffness-related features in both controlled and real-world conditions.

Expected Outcome

- A suite of signal processing and machine learning tools for analysing wrist pulse waveforms and estimating arterial stiffness.
- Visualizations and statistical analyses that clarify how arterial stiffness alters pulse waveform morphology.
- Models validated on both synthetic data and preliminary real PPG data collected from the wrist
- Documentation and code repository for the workflow, supporting reproducibility and further development.

Learning Goals

By the end of this project, I aim to:

1. Broaden my medical and clinical knowledge, and gain insight into the world of health technology and digital health solutions.
2. Develop practical skills in signal processing, such as noise filtering, segmentation, and feature engineering, specifically tailored to physiological waveform data.
3. Gain hands-on experience with machine learning for biomedical signals, including model training, validation, and performance interpretation.
4. Demonstrate the transferability of methods—by successfully applying models and insights from synthetic data to real PPG data acquired from wearable devices.
5. Produce clear visualizations and documentation to communicate technical results and support future research or development.

Tasks / Work Breakdown Structure (WBS)

1. Data Preparation & Exploration (20 points)

1.1. Downloading and understanding synthetic datasets (5)
1.2. Data loading, cleaning, plausibility filtering (5)
1.3. Signal processing: filtering, truncation, normalization, sliding window (5)
1.4. Initial visualization and statistical exploration (5)

2. Feature Engineering & Classical Analysis (20 points)

2.1. Extracting classical PPG indices (RI, SI, AGI mod) (2)
2.2. Calculating additional morphological features (8)
2.3. Correlation analysis with PWV and visualization (7)
2.4. Documentation of feature extraction pipeline (3)

3. Machine Learning Modelling (20 points)

3.1. Feature-based regression (using classical features) (4)
3.2. Waveform-based regression (using full signals) (4)
3.3. Model selection, hyperparameter tuning, and validation (4)
3.4. Interpretation of model results and feature importance (4)
3.5. Benchmarking ML performance against classical methods (4)

4. Integration with Real Data (20 points)

4.1. Real-world wrist PPG data collection (if available) (5)
4.2. Preprocessing and quality assessment of real data (5)
4.3. Application and adaptation of models to real data (6)
4.4. Comparative analysis: synthetic vs. real results (4)

5. Visualization & Communication (10 points)

5.1. Creating clear, insightful plots and figures (5)
5.2. Writing documentation, README, and user instructions (3)
5.3. Preparation of presentation materials or demo (2)

6. Clinical & Health Technology Understanding (10 points)

6.1. Literature review on arterial stiffness and wearable sensing (5)
6.2. Mapping findings to clinical relevance and future health tech applications (5)

Summary Table

Task Area	**Points**
1. Data Preparation & Exploration	20
2. Feature Engineering & Classical Analysis	20
3. Machine Learning Modelling	20
4. Integration with Real Data	20
5. Visualization & Communication	10
6. Clinical & Health Technology	10
Total	**100**

Schedule and key milestones

Weeks	Task/Phase	Description	Milestone
1-2	Loop 1: Initial Pipeline	Data prep, visualization, signal processing, ML model on synthetic data	M1: First pipeline ready, 1st supervisor meeting **(Friday, 4th July 2025**)
3-5	Loop 2: Refinement Cycle	Add/focus features, adjust data or models per supervisor input, rerun analysis	M2: Enhanced features/models, supervisor check-in (**Friday, 25th July 2025)**
6-8	Loop 3: Further Iteration	Deeper or focused analysis, possibly explore additional features, robustness, or more data sources	M3: Analysis matured, finalize synthetic results **(Friday, 15th August 2025)**
9-12	Real Data Application	Collect real PPG data, preprocessing, apply models, analyse real-vs-synthetic, document results	M4: Real data results, project wrap-up **(Mid of September 2025**)

Critical Path/Order:

- Loop 1 → Supervisor feedback → Loop 2 → Supervisor feedback → Loop 3 → Supervisor feedback → Real Data Application → Feedback → Wrap-up

Identified Risks

Risk	Mitigation Plan
1. Real wrist PPG data collection is delayed, or hardware is faulty	Continue improving methods on synthetic data; prepare simulated "noisy" signals; schedule backup device testing.
2. Machine learning model fails to generalize from synthetic to real	Test robustness with added noise, retrain with small real datasets, or try transfer learning/adaptation methods.
3. Insufficient or poor-quality real data	Plan multiple recording sessions, implement strong signal quality checks, and document data limitations clearly.
4. Model performance is unsatisfactory (low accuracy, unstable)	Revisit feature engineering, try alternative models, or consult literature/supervisor for new approaches.

Project Management and Coordination

- Progress Tracking: Progress will be tracked using a combination of weekly self-checks, GitHub code commitments, and bi-weekly supervisor meetings.
- Milestone Attainment: Key milestones (pipeline complete, supervisor reviews, synthetic results, real data integration) are clearly marked in the schedule. After each, results will be summarized and evaluated with the supervisor.
- Change Management: If a major change or risk occurs (e.g., real data unavailable), this will be discussed in the next supervisor meeting, and schedule will be updated to reflect revised priorities and tasks.

Documentation

Format:
Project will be documented using a GitHub repository containing:
- A running lab/project log (decisions, experiments, problems encountered)
- Well-commented code/scripts
- A README file and user guide for data and scripts
- Visualizations and summary figures
- Final project report and presentation slides

Deadlines:
Key documents (e.g., README, code, logbook) will be updated at the end of each iteration loop (every 2–3 weeks), and the final report will be completed and submitted at the project deadline.

Project Success Indicators

1. Reaching the goals: Success means that the project achieves the stated learning goals: understanding arterial stiffness encoding in pulse waveforms, building effective analysis/modelling tools, and applying them to both synthetic and real data.
2. Meeting deadlines: Each milestone in the Gantt chart will be used to track on-time progress.
3. Commitment: Demonstrated by consistent work, regular communication, supervisor meetings, and detailed project logs.
4. Quality of documentation: Complete, clearly written README, well-commented code, lab book, and a final report with clear results and interpretation.
5. Quality of design: Systematic signal processing and ML workflow, justified design choices, and well-organized code and visualizations.
6. Learning: Concrete evidence of new knowledge in signal processing, machine learning, medical/clinical concepts, and digital health; reflections included in the final report.

Plan to Meet These:
Follow the planned workflow and regularly update documentation.
Maintain clear and realistic timelines with milestones.
Solicit and act on supervisor and peer feedback.
Submit all deliverables on time and review each for completeness and clarity.