COMILLAS
UNIVERSIDAD PONTIFICIA
ICAI

GRADO EN INGENIERÍA EN TECNOLOGÍAS DE TELECOMUNICACIÓN

TRABAJO FIN DE GRADO

SIGNAL ANALYSIS FOR MONITORING AN EXTRACORPOREAL MEMBRANE OXYGENATION SYSTEM IN REAL TIME

Autor: Clara Palacios Castrillo

Supervisor: Lu Li

Co-Supervisor: Brittany Gómez

Madrid

I hereby declare, under my own responsibility, that the Project entitled:

Signal Analysis for Monitoring an Extracorporeal Membrane Oxygenation System in Real Time

submitted at the ICAI School of Engineering of Comillas Pontifical University during the 2024/25 academic year, is my own work, original and unpublished, and has not been previously submitted for any other purpose.

This Project is not plagiarized from any another work, either in whole or

in part, and any information taken from other sources is properly cited.

Clara Palacios

Signed by: Clara Palacios Castrillo Date: 19./ June/ 2025

Project submission authorized by

PROJECT SUPERVISORS

Signed by: Lu Li Date: 06./ 19./ 2025

Signed by: Brittany Gómez Date: 06./ 20./ 2025 Brittany S. Gomez

COMILLAS
UNIVERSIDAD PONTIFICIA
ICAI

GRADO EN INGENIERÍA EN TECNOLOGÍAS DE TELECOMUNICACIÓN

TRABAJO FIN DE GRADO

SIGNAL ANALYSIS FOR MONITORING AN EXTRACORPOREAL MEMBRANE OXYGENATION SYSTEM IN REAL TIME

Autor: Clara Palacios Castrillo

Supervisor: Lu Li

Co-Supervisor: Brittany Gómez

Madrid

Acknowledgments

Thank you to my family for their unwavering support and belief in me throughout this academic journey.

To my friends, I am grateful for your companionship, encouragement, and the laughter that made everything more bearable.

I am also thankful to my supervisors for their dedicated support, thoughtful insights, and constructive feedback.

ANÁLISIS DE SEÑALES PARA MONITORIZAR EN TIEMPO REAL UN SISTEMA DE OXIGENACIÓN EXTRACORPOREA POR MEMBRANA

Autor: Palacios Castrillo, Clara.
Director: Li, Lu y Gomez, Brittany S.
Entidad Colaboradora: Carnegie Mellon University

RESUMEN DEL PROYECTO

Esta tesis desarrolla un sistema basado en el active learning para la predicción de fallos en tiempo real y la monitorización de máquinas ECMO (*Extracorporeal Membrane Oxygenation*) portátiles, con el objetivo de detectar anomalías como la formación de coágulos de sangre y alertar a los médicos. El sistema también estima la frecuencia respiratoria y el ritmo cardiaco del paciente con señales no invasivas como auriculares de botón. Utilizando datos in-vitro e in-vivo demostramos una detección eficaz de anomalías y una precisión prometedora para la monitorización de las constantes vitales.

Palabras clave: ECMO, *Active Learning*, Detección de Fallos, Ritmo Respiratorio, Salud cardiaca, Dispositivos médicos, *Machine Learning*, *Portable Healthcare*

1. Introducción

Los dispositivos médicos monitorizados a distancia están transformando la asistencia sanitaria a domicilio al reducir las estancias hospitalarias y proporcionar comodidad a los pacientes que esperan tratamientos críticos. Los sistemas ECMO (*Extracorporeal Membrane Oxygenation*) son esenciales para pacientes con insuficiencia pulmonar grave en espera de un trasplante, pero su uso hospitalario prolongado puede provocar complicaciones como atrofia muscular (van Diepen et al., 2021). Este problema se evita utilizando sistemas ECMO portátiles que facilitan el movimiento y permiten su uso fuera del hospital. La detección eficaz de fallos en los sistemas ECMO portátiles es esencial para mejorar la seguridad del paciente, permitiéndole vivir en casa. Esta investigación presenta un sistema basado en *machine learning* que monitoriza el rendimiento de los sistemas ECMO y las constantes vitales del paciente.

2. Definición del proyecto

El objetivo principal del proyecto es diseñar un sistema que pueda monitorizar eficazmente el estado de una máquina ECMO portátil en tiempo real, identificando posibles fallos como la formación de coágulos de sangre, mediante sensores de presión y flujo en la máquina. Además, este sistema de monitorización utiliza el sensor de presión para estimar la frecuencia respiratoria, y auriculares para reconstruir las señales cardiacas del

sismocardiograma (SCG). Mediante técnicas de *machine learning*, el sistema se adapta a las variaciones de los datos del paciente, minimizando la necesidad de extensos conjuntos de datos etiquetados.

3. Descripción del sistema

El sistema desarrollado en este proyecto se centra en la monitorización del funcionamiento de la máquina ECMO portátil, la detección de la frecuencia respiratoria y el seguimiento de la salud cardiovascular. Cada uno de estos componentes se basa en diferentes señales fisiológicas, que se procesan mediante modelos de *machine learning* para realizar predicciones en tiempo real.

El sistema de monitorización de ECMO utiliza datos de sensores de presión y flujo en tiempo real para detectar anomalías como coágulos de sangre u obstrucciones en los catéteres intravenosos. Los modelos de *machine learning* son capaces de identificar patrones anormales en estas señales, y *active learning* garantiza que el modelo se adapte a los nuevos datos con eficacia, sin necesidad de extensos conjuntos de datos etiquetados. Este sistema se ha desarrollado a partir de datos in-vitro y se ha validado con datos in-vivo.

Para la monitorización respiratoria, el sistema detecta patrones en los datos de presión que se corresponden con la frecuencia respiratoria del paciente. Esta información adicional proporciona un contexto que permite evaluar correctamente el estado del paciente, distinguiendo entre los cambios debidos a los movimientos o la actividad física del paciente y los posibles fallos del sistema ECMO.

La monitorización cardíaca se consigue utilizando auriculares equipados con micrófonos para captar señales de fonocardiograma (PCG). Estas señales se transforman posteriormente a señales de sismocardiograma (SCG) mediante un modelo de red neuronal. Este enfoque no invasivo ofrece un método accesible y continuo para controlar la salud cardiovascular, y las redes neuronales de tipo transformer han demostrado ser especialmente eficaces para esta tarea.

4. Resultados

Los modelos de *machine learning* desarrollados para la **detección de fallos de ECMO** se probaron con datos in-vitro e in-vivo. El sistema detecta correctamente anomalías como caídas de presión lentas y repentinas en los circuitos de ECMO, y el modelo de *active learning* se adaptó de manera muy eficiente a los nuevos datos. Entre las estrategias de consulta probadas, «Stream Probabilistic Active Learning» demostró el mejor rendimiento, manteniendo una alta precisión durante los cambios de presión y garantizando la adaptabilidad del sistema a las condiciones variables del paciente, al tiempo que tenía un

recuento de adquisiciones razonable. El modelo también se validó utilizando datos in-vivo de ovejas (Figura 1). donde detectó eventos de succión y se ajustó a los cambios en la posición de la oveja, distinguiendo entre el comportamiento normal y los fallos potenciales. Esta adaptabilidad es crucial para entornos clínicos en los que el movimiento del paciente u otros factores ambientales pueden afectar a los datos.

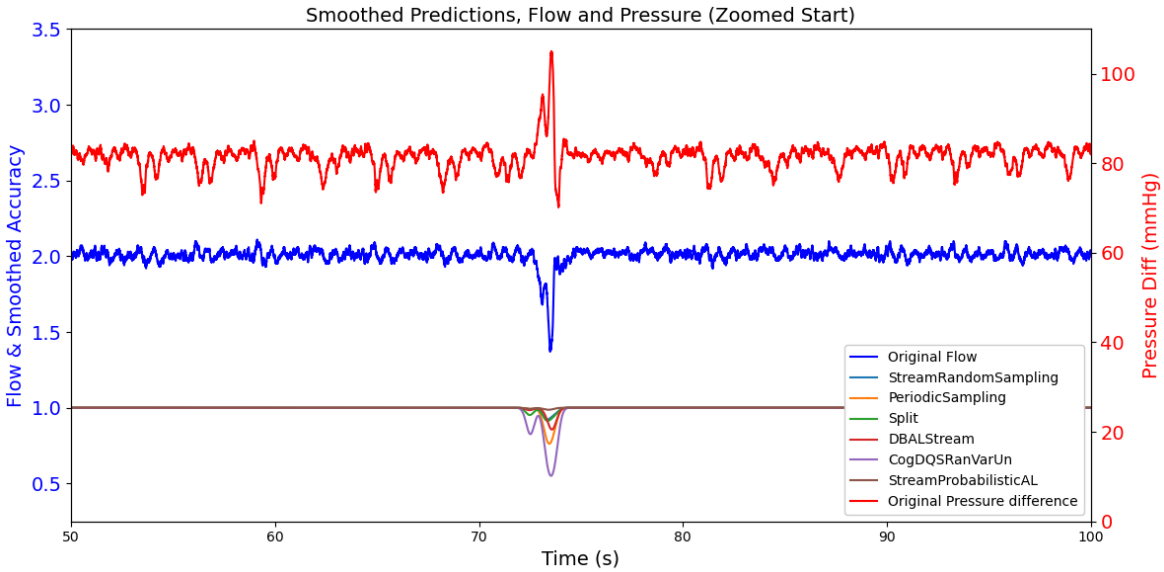

Figura 1. Predicciones de todos los métodos de consulta sobre un evento de succión (gráfico inferior) y mostrando el caudal (L/min) en azul y la presión en rojo

Para **estimar la frecuencia respiratoria**, se utilizaron métodos de detección de picos en el dominio temporal y de transformada de Fourier (FFT). El método de detección de picos produjo una estimación precisa de alrededor de 6,5 respiraciones por minuto, consistente con la frecuencia respiratoria esperada de la oveja (Abrams et al., 2020), como se muestra en la Figura 2. El método FFT también proporcionó resultados fiables, aunque carecía de la precisión del método en el dominio temporal (Skotte y Kristiansen, 2014).

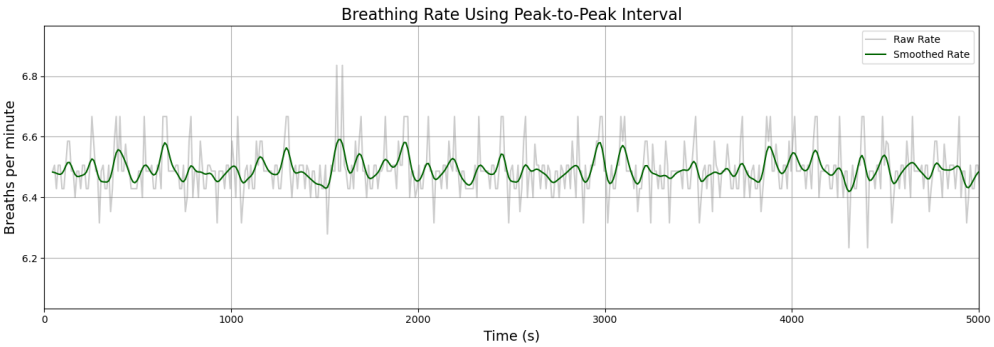

Figura 2. Frecuencia respiratoria suavizada a lo largo del experimento

El modelo de **monitorización de la salud cardiovascular**, que transforma las señales PCG en señales SCG, se implementó con éxito. El modelo basado en redes neuronales de tipo transformer superó tanto a las redes neuronales convolucionales como a las recurrentes en la conversión de PCG a SCG, como se muestra en la Figura 3. El sistema demostró su

potencial para la monitorización cardiaca no invasiva mediante el uso de auriculares normales, proporcionando una alternativa más sencilla a los métodos tradicionales de ECG (Figura 4). A pesar de algunos problemas con la alineación de la señal, el modelo de redes neuronales de tipo transformer captó la estructura general de la forma de onda SCG, lo que lo convierte en una solución prometedora para la monitorización cardiaca continua en casa.

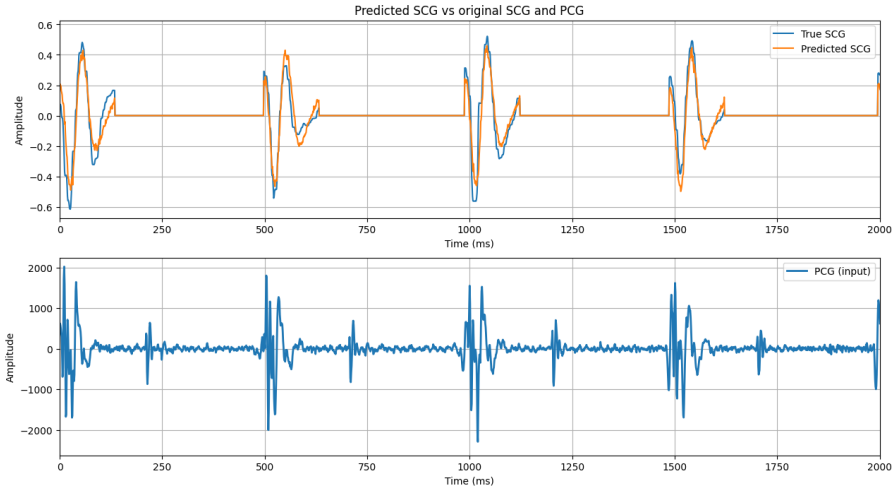

Figura 3. Salida de la redes neuronales de tipo transformer de la señal SCG del PCG captada con un teléfono

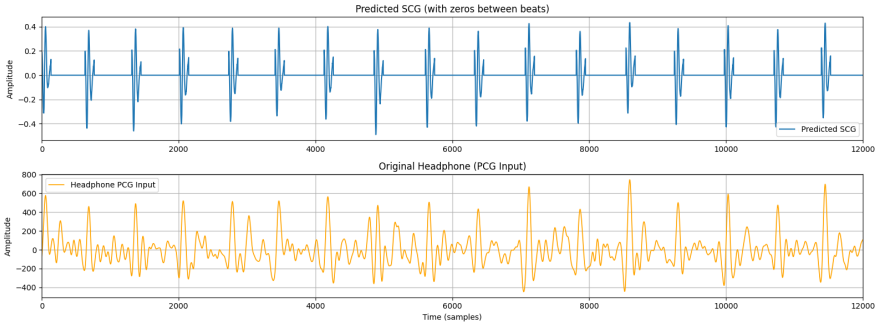

Figura 4. Recreación de la señal cardíaca SCG de PCG captada con auriculares

5. Conclusión

Esta investigación desarrolló con éxito un sistema basado en *active learning* para la monitorización en tiempo real de dispositivos ECMO portátiles, capaz de detectar fallos críticos. Las técnicas de *active learning* permitieron al sistema afrontar con eficacia el reto de los datos etiquetados limitados, garantizando la adaptabilidad a diferentes pacientes y condiciones. Además, se desarrollaron modelos para monitorizar las constantes vitales del paciente. Al utilizar sensores portátiles como auriculares, el sistema ofrece una forma no invasiva y rentable de monitorizar el corazón. El trabajo futuro se centrará en seguir optimizando el rendimiento del sistema, integrar sensores adicionales y ampliar las pruebas para incluir diversos escenarios clínicos.

6. Referencias

Abrams, D., Schmidt, M., Pham, T., Beitler, J. R., Fan, E., Goligher, E. C., McNamee, J. J., Patroniti, N., Wilcox, M. E., Combes, A., Ferguson, N. D., McAuley, D. F., Pesenti, A., Quintel,

M., Fraser, J., Hodgson, C. L., Hough, C. L., Mercat, A., Mueller, T., … Brodie, D. (2020). Mechanical ventilation for acute respiratory distress syndrome during extracorporeal life support research and practice. *American Journal of Respiratory and Critical Care Medicine*, 201(5), 514–525. https://doi.org/10.1164/RCCM.201907-1283CI

Skotte, J. H., & Kristiansen, J. (2014). Heart rate variability analysis using robust period detection. *BioMedical Engineering Online*, 13(1), 1–11. https://doi.org/10.1186/1475-925X-13-138

van Diepen, A., Bakkes, T. H. G. F., De Bie, A. J. R., Turco, S., Bouwman, R. A., Woerlee, P. H., & Mischi, M. (2021). A Model-Based Approach to Synthetic Data Set Generation for Patient-Ventilator Waveforms for Machine Learning and Educational Use. *Journal of Clinical Monitoring and Computing*, 36(6), 1739–1752. https://doi.org/10.1007/s10877-022-00822-4

SIGNAL ANALYSIS FOR MONITORING AN EXTRACORPOREAL MEMBRANE OXYGENATION SYSTEM IN REAL TIME

Author: Palacios Castrillo, Clara.
Supervisor: Li, Lu and Brittany Gomez
Collaborating Entity: Carnegie Mellon University

ABSTRACT

This thesis develops an active learning-based system for real-time failure prediction and monitoring of portable ECMO (Extracorporeal Membrane Oxygenation) machines, aiming to detect anomalies like blood clot formation and alert doctors. The system also estimates the patient's breathing rate and heart rhythm with non-invasive signals such as in-ear headphones. Using in-vitro and in-vivo data we demonstrate effective anomaly detection and promising accuracy for vitals monitoring.

Keywords: ECMO, Active Learning, Failure Detection, Breathing Rate, Heart Health, Medical Devices, Machine Learning, Portable Healthcare

1. Introduction

Remotely monitored medical devices are transforming home healthcare by reducing hospital stays and providing comfort to patients awaiting critical treatments. ECMO (Extracorporeal Membrane Oxygenation) systems are essential for patients with severe lung failure awaiting transplants, but long-term use can lead to complications like muscle atrophy (van Diepen et al., 2021). Portable ECMO devices that encourage patient movement could solve this problem, even allowing them to leave the hospital. Effective failure detection in portable ECMO systems could significantly enhance patient safety, granting the patient the possibility to live at home. This research presents a machine learning-based system that monitors ECMO performance and vitals of the patient.

2. Project Definition

The primary objective of the project is to design a system that can effectively monitor the condition of a portable ECMO machine in real time, identifying potential failures such as blood clot formation, through pressure and flow sensors in the machine. Additionally, this monitoring system utilizes the pressure sensor to estimate breathing rate, and in-ear headphones to reconstruct seismocardiogram heart signals (SCG). Through active learning techniques, the system adapts to variations in patient data, minimizing the need for extensive labeled datasets.

3. System Description

The system developed in this project focuses on monitoring the portable ECMO machine's performance, detecting breathing rate, and tracking heart health. Each of these components relies on different physiological signals, which are processed through machine learning models for real-time predictions.

The ECMO monitoring system utilizes real-time pressure and flow sensor data to detect anomalies such as blood clots or tubing obstructions. Machine learning models are capable of identifying abnormal patterns in these signals, and active learning ensures the model adapts to new data efficiently, without requiring extensive labeled datasets. This system was developed using in-vitro data and then validated on in-vivo data.

For respiratory monitoring, the system detects patterns in the pressure data that correspond to the patient's breathing rate. This additional information provides context for evaluating the patient's condition, distinguishing between changes due to the patient's movements or physical activity and potential failures of the ECMO system.

Heart monitoring is achieved using in-ear headphones equipped with microphones to capture phonocardiogram (PCG) signals, which are then mapped to seismocardiogram (SCG) signals using a neural network model. This non-invasive approach offers an accessible and continuous method for monitoring heart health, with transformers proving to be particularly effective for this task.

4. Results

The machine learning models developed for **ECMO failure detection** were tested on both in-vitro and in-vivo data. The system successfully detected anomalies such as slow and sudden pressure drops in ECMO circuits, with the active learning model adapting well to new data. Among the query strategies tested, "Stream Probabilistic Active Learning" demonstrated the best performance, maintaining high accuracy during pressure changes and ensuring the system's adaptability to varying patient conditions, while also having a reasonable acquisition count. The model was also validated using in-vivo sheep data (Figure 1), where it detected suction events and adjusted to changes in the sheep's position, distinguishing between normal behavior and potential failures. This adaptability is crucial for clinical settings where the patient's movement or other environmental factors can affect the data.

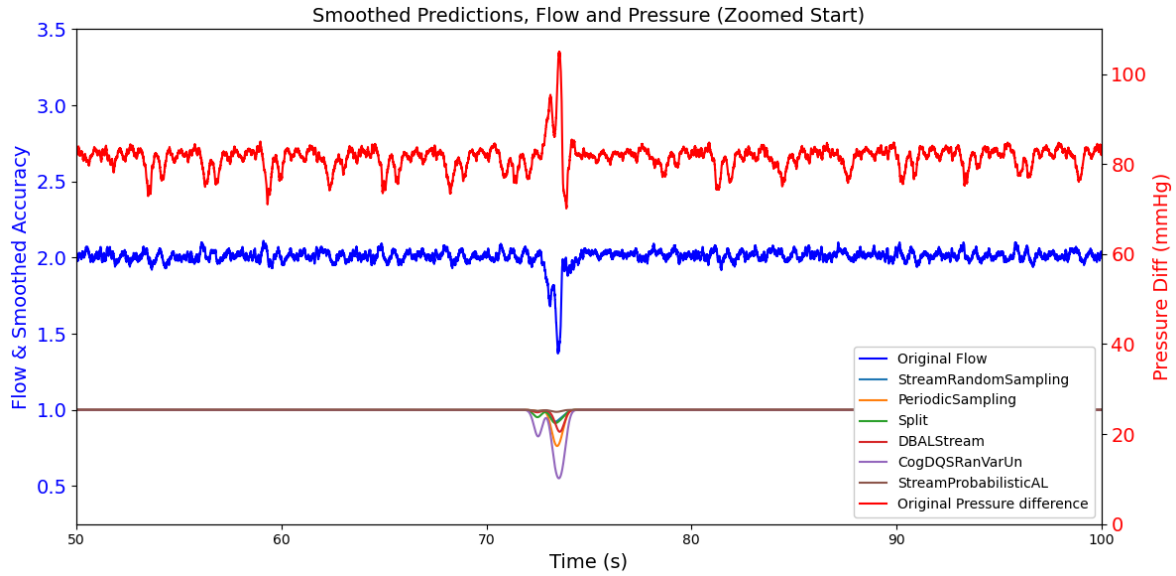

Figure 1. Predictions of all query methods on a suction event (bottom plot) and showing the flow rate (L/min) in blue and the pressure in red

For **breathing rate estimation**, both peak detection in the time domain and Fourier transform (FFT) methods were used. The peak detection method produced an accurate estimate of around 6.5 breaths per minute, consistent with the sheep's expected breathing rate (Abrams et al., 2020), as shown in Figure 2. The FFT method also provided reliable results, though it lacked the precision of the time-domain method (Skotte & Kristiansen, 2014).

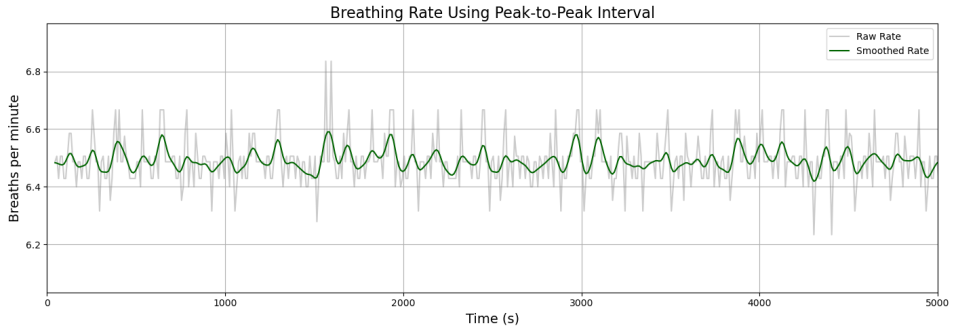

Figure 2. Smoothed breathing rate over the experiment

The **heart health monitoring** model, which transforms PCG signals into SCG signals, was successfully implemented. The transformer-based model outperformed both convolutional and recurrent neural networks in mapping PCG to SCG, shown in Figure 3. The system showed the potential for non-invasive heart monitoring using regular in-ear headphones, providing a simpler alternative to traditional ECG methods (Figure 4). Despite some challenges with signal alignment, the transformer model captured the SCG waveform's overall structure, making it a promising solution for continuous heart monitoring at home.

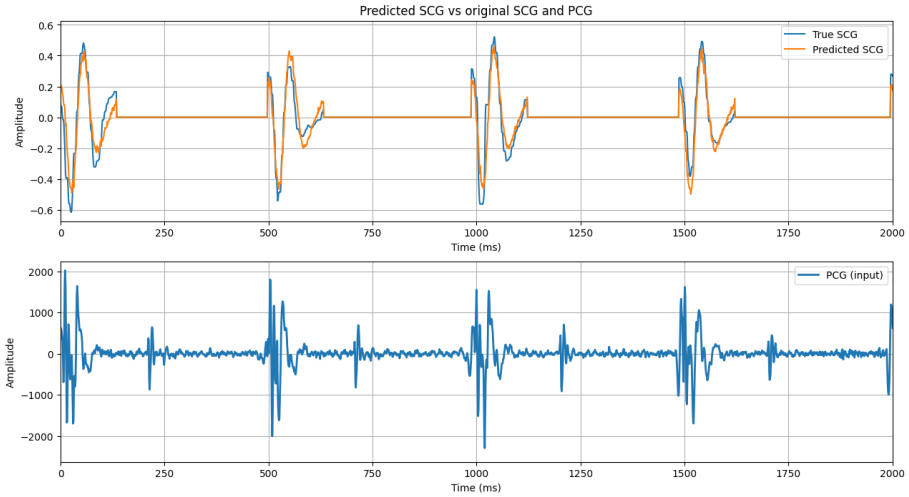

Figure 3. Transformer's output of SCG signal from PCG captured with a phone

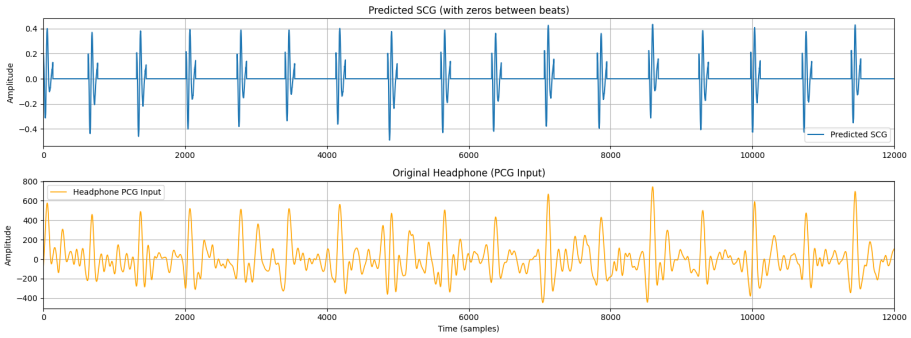

Figure 4. Recreation of SCG heart signal from PCG captured from headphones

5. Conclusion

This research successfully developed an active learning-based system for real-time monitoring of portable ECMO devices, capable of detecting critical failures. Active learning techniques allowed the system to efficiently handle the challenge of limited labeled data, ensuring adaptability to different patients and conditions. In addition, models were developed to monitor patient's vitals. By utilizing wearable sensors like in-ear headphones, the system provides a non-invasive, cost-effective way to monitor heart. Future work will focus on further optimizing the system's performance, integrating additional sensors, and expanding testing to include diverse clinical scenarios.

6. References

Abrams, D., Schmidt, M., Pham, T., Beitler, J. R., Fan, E., Goligher, E. C., McNamee, J. J., Patroniti, N., Wilcox, M. E., Combes, A., Ferguson, N. D., McAuley, D. F., Pesenti, A., Quintel, M., Fraser, J., Hodgson, C. L., Hough, C. L., Mercat, A., Mueller, T., Brodie, D. (2020). Mechanical ventilation for acute respiratory distress syndrome during extracorporeal life support research and practice. *American Journal of Respiratory and Critical Care Medicine*, 201(5), 514–525. https://doi.org/10.1164/RCCM.201907-1283CI

Skotte, J. H., & Kristiansen, J. (2014). Heart rate variability analysis using robust period detection. *BioMedical Engineering Online*, 13(1), 1–11. https://doi.org/10.1186/1475-925X-13-138

van Diepen, A., Bakkes, T. H. G. F., De Bie, A. J. R., Turco, S., Bouwman, R. A., Woerlee, P. H., & Mischi, M. (2021). A Model-Based Approach to Synthetic Data Set Generation for Patient-Ventilator Waveforms for Machine Learning and Educational Use. *Journal of Clinical Monitoring and Computing*, 36(6), 1739–1752. https://doi.org/10.1007/s10877-022-00822-4

Table of Contents

List of Figures

List of Tables

Chapter 1. INTRODUCTION

Remotely monitored medical devices are the next stage in quality-of-life home care, leaving behind the restless hospital stays and capitalizing on the comfort of home. Currently, patients that have an injury to their lungs and need a transplant can be on a ventilator for up to a year in a hospital until the transplant becomes available (Valapour et al., 2024). Extracorporeal Membrane Oxygenation (ECMO) is a system that extracts blood from the patient for oxygenation and returns it to the body, working as an external lung. Long use of ECMO and previous generations of ventilation are both associated with muscle atrophy and weakness in patients (van Diepen et al., 2021). There is evidence that ECMO can attenuate some of this loss when participating in physical activity, like ambulation. However, only an estimated 35% of ECMO patients were able to exercise and only 61% of those patients could walk, which has been shown to fight muscle atrophy (Abrams et al., 2022; Hayes et al., 2018).

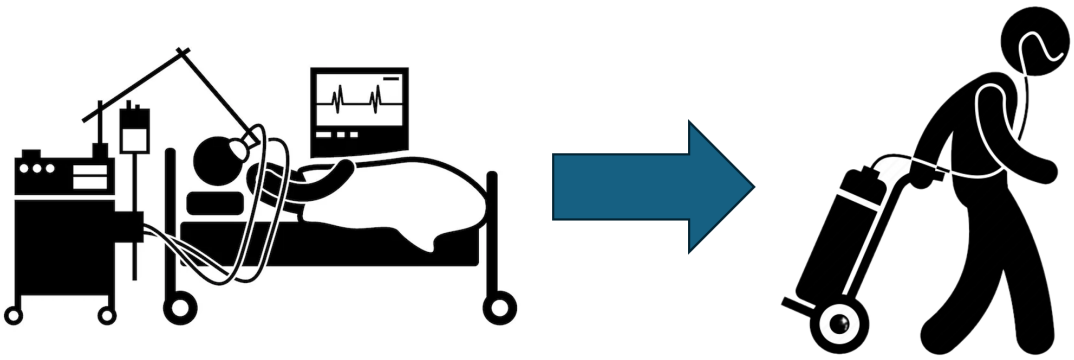

Figure 1. Portable ECMO machines allow patients to stay at home and have a better quality of life as they wait for a transplant

A smaller, lighter and portable redesigned ECMO prototype could provide a patient with the comfort of being at home and encourage ambulating, while maintaining a high standard of care (Figure 1). However, this device would require complex clinical and life-threatening

events to be detected, along with a notification system to the patient and clinician with enough time to address the issue. This thesis focuses on the analysis of the signals collected from the portable ECMO device to ensure it is working properly and guarantee the health of the patient.

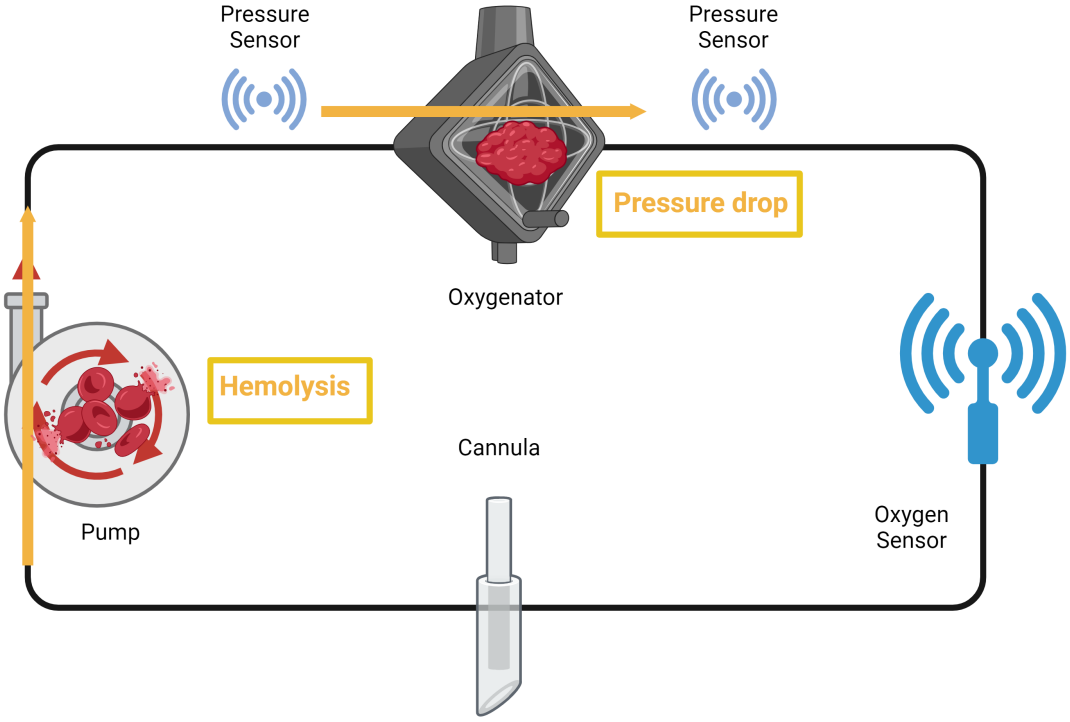

Figure 2. Circuit of an artificial lung along with typical complications and where they can occur in the circuit (Gomez et al., 2025)

Of the events of interest that should be monitored in case of failure, the main one is blood clot formation which causes reduced oxygen transfer rate. The main objective of this work is to develop an automatic classifier to determine if the ECMO system is operating properly and the patients' vitals are in a healthy range, based on signals collected in real-time and though the patient's in-ear headphones. There is a challenge associated with detecting failure cases, in addition to the use of medical devices, by the fact that there will be physiological

differences between in-silico, in-vitro, in-vivo, and even data collected through clinical experiments with a prototype medical device. Therefore, to build a malleable classifier that can work on little data, active learning (AL) techniques are proposed to detect and predict clot formation and monitor breathing rate. Finally, to monitor heart rate, signal is recorded through in-ear headphones and processed using transformers and convolution neural networks to obtain an accurate seism cardiogram (SCG) heart signal.

Chapter 2. DESCRIPTIONS OF THE TECHNOLOGIES

The main goal of this project is to develop a system that can detect physiological anomalies in real time using biological signals. Due to the large volume and variability of biomedical data, particularly unlabeled recordings, we utilize machine learning techniques that can adapt to new patients and data distributions. This section describes the core technologies that enable this adaptability. These include active learning for efficient data annotation and neural network architectures suited to time series and spectral analysis.

Machine learning is a branch of artificial intelligence that enables computers to learn patterns from data and make predictions or decisions without being explicitly programmed. Supervised learning trains models on labeled data to predict known outcomes, while unsupervised learning finds hidden patterns in unlabeled data. Active learning is a supervised approach where the model selectively queries for labels on the most informative data points to improve performance with fewer labeled samples. However, some applications require the processing of large amounts of unlabeled data, such as in medicine. In this field, data is often collected at such a large scale that manual labeling becomes infeasible. In the medical field, this is due to the need for expert annotations, which are time-consuming, costly, and often limited by clinicians' variability —where different clinicians may label the same data inconsistently. Additionally, the complexity and volume of medical data make large-scale manual labeling impractical. This challenge is addressed by **Active Learning** (Settles, 2009). Active learning algorithms select the most informative data samples from which to learn. Information theory provides a mathematical framework for quantifying uncertainty and information gain, which is essential in guiding learning algorithms toward the most informative data. Active learning queries an oracle for labels of those samples expected to most improve the model. This results in improved performance using fewer labeled examples. Active Learning aims to achieve high accuracy using as few labeled instances as possible, minimizing the cost of obtaining labeled data.

A particular case of active learning is called pool-based active learning (Settles, 2009). In this scenario, the model selects the most informative subsets of observations from a set of unlabeled data. Online Active Learning is designed for data streams (Settles, 2009), where data arrives continuously. Instances are generated in a continuous stream and cannot be stored in their entirety before a decision is made. Each incoming data point is observed and evaluated in real time. Then, it is determined if it is informative enough, or should be discarded. This approach was used in the model as it presented two main advantages. The first is that it allows the model to process data as it comes in and make predictions in real time, adapting to different situations, such as walking, sleeping, going up stairs, etc. This is because the algorithm can ask about changes in the data and depending on the label, use it to adapt the model to a new 'normal situation'. The second advantage is its ability to adapt on to different patients. By selectively querying informative examples, the model can adapt thresholds to better match each patient's physiological baseline, improving anomaly detection.

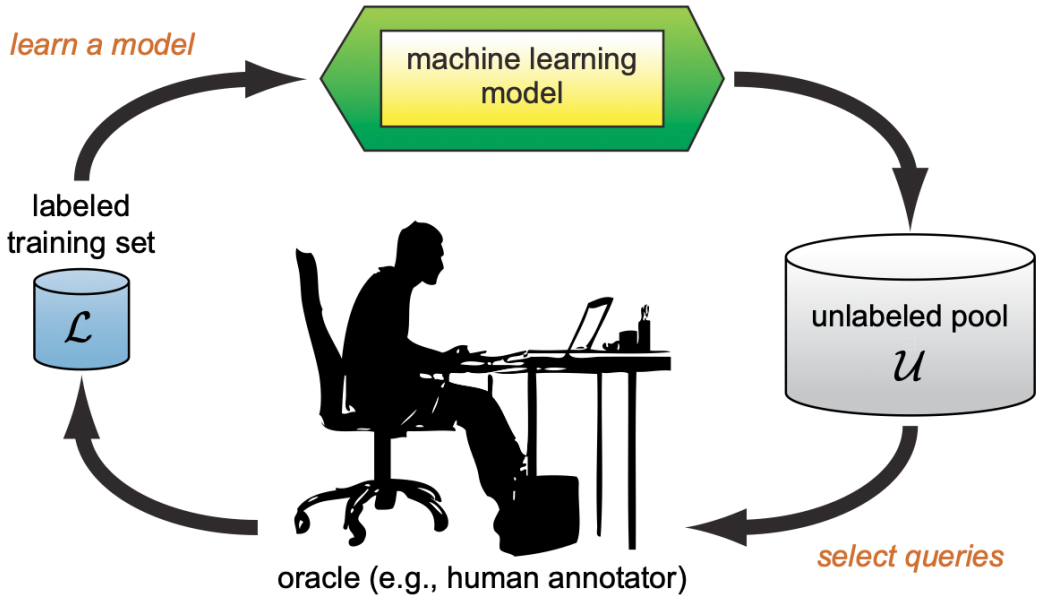

Figure 3. The pool-based active learning cycle.(Settles, 2009)

For heartbeat detection and for converting from one type of heart signal to another, **neural network models** are well suited as they are able to identify and learn more complex patterns in the data. Neural networks are inspired from the structure of the human brain, more specifically our neurons and the connections or synapses between them. They are composed of different layers which are divided into 3 groups: input layers, hidden layers and output layers. The input layer receives the raw data and hands it off to the hidden layers that process information and finally it ends in the output layer which produces a final prediction or classification. Layers are composed of neurons which compute a weighted sum of inputs and then a non-linear activation function (Nielsen, 2015a). Neural networks learn patterns by adjusting weights and biases in the neurons by using backpropagation and gradient descent. Backpropagation is a method of computing the gradients of the loss function with respect to each weight in the network. It is based on the calculus chain rule and works by performing a forward pass to compute prediction and loss, followed by a backwards pass that moves from output to input layers and calculates how much each weight contributed to the error. Gradient descent (Nielsen, 2015a) is an optimization algorithm used to minimize the loss function. It uses gradients to update the weights:

$$w_{\text{new}} = w_{\text{new}} - \eta \frac{\delta Loss}{\delta w}$$

where η is the learning rate. This updates the weights in the direction that reduces error. There are several variants that differ in update frequency and convergence properties, such as batch gradient descent, stochastic gradient descent, mini-batch gradient descent or Adam, RMSdrop, etc. This process of adjusting weights and biases in the neurons by using backpropagation and gradient descent, allows each neuron to know how to adjust its weights to reduce error. They can learn complex functions and decision boundaries and usually require large datasets to train, which means high computational power and are often trained on GPUs.

There are different types of neural networks depending on the depth and use. Convolutional neural networks (CNN) are designed to process grid-like data and use convolutional filters to detect local patterns, such as edges and textures in images. In biomedical signal

processing, this architecture is particularly effective when signals are transformed into time-frequency representations—for example, spectrograms or short-time Fourier transforms (STFT)—which resemble images. These two-dimensional matrices capture both temporal and frequency information, allowing CNNs to identify characteristic signal patterns, such as those associated with heartbeat morphology or epileptic discharges. Hidden layers typically include convolution, activation, pooling and fully connected layers. The convolution layers apply filters or kernels to detect local patterns, like edges. The filter slides across the data and performs element wise convolution and produces a feature map to highlight specific features. This layer reduces the number of parameters compared to fully connected layers. Activations layers add non-linearity to the model which allows it to learn complex patterns. The most common ones are ReLu, Sigmoid and Tanh. A pooling layer reduces spatial dimensions of feature maps and helps to down sample and reduce computational complexity. Some common ones are max pooling, and average pooling. This process makes the features more robust to translations and noise. Finally, fully connected layers mean that every neuron is connected to all neurons in the previous layer. They are typically used near the end of the network for classification or regression. It converts the high-level learned features into final outputs. It's great at extracting spatial hierarchies in data (low level features and high-level objects). They are most commonly used in image classification, object detection and video analysis.

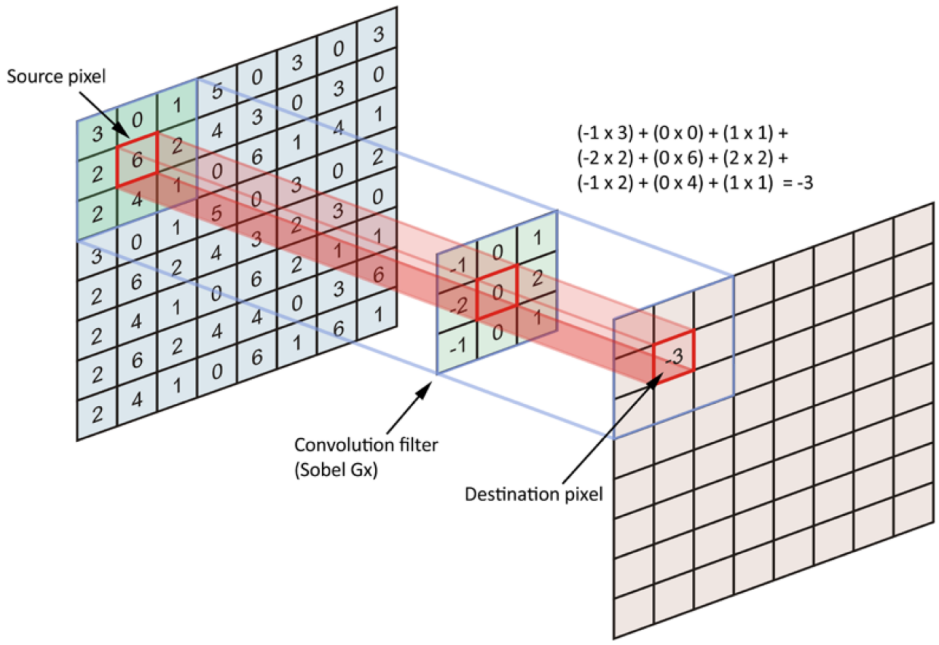

Figure 4. Convolution layer diagram (Nielsen, 2015b)

Recurrent neural networks (RNN) are designed to process sequential data, such as time series, speech or text. This model maintains a hidden state that captures information about previous time steps. This internal memory is updated at each time step:

$$h_t = f(h_{t-1}, x_t)$$

This allows the model to remember context in sequences and is essential for time series predictions, similar to a short-term memory within the model (Figure 5).

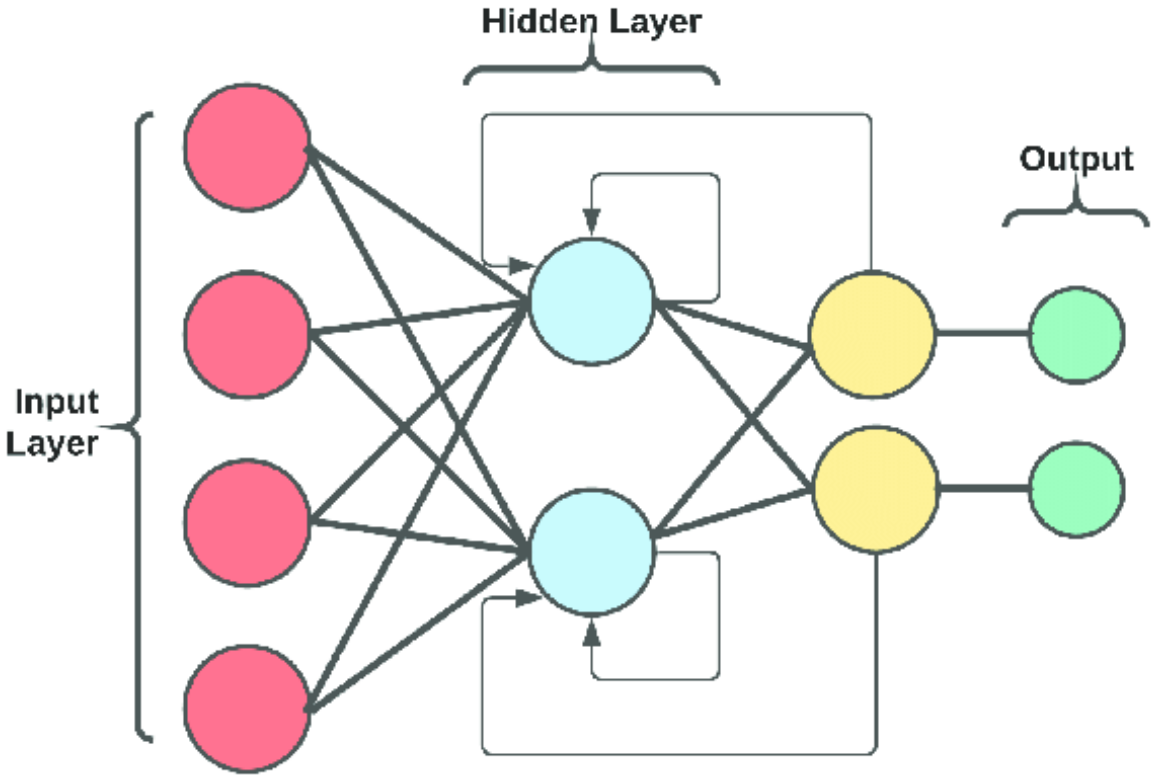

Figure 5. Recurrent Neural Network diagram (Arias et al., 2022)

In this case, each output depends not just on the current input but also on past inputs and outputs. RNNs struggle with long-term dependencies due to vanishing gradient problems. This is when the gradients become close to zero as they're passed back though the many layers, which means that early layers learn slower or nearly zero compared to the ones towards the output. They are commonly used in time series forecasting, speech recognition, and also natural language processing (NLP).

Transformers are designed to handle long sequential data inputs without recurrent connections or hidden states, unlike RNNs. It uses self-attention mechanisms to weigh the importance of different input elements (Vaswani et al., 2017). The main components are multi-head self-attention (Figure 6), which allows the model to focus on multiple parts of the sequence simultaneously, positional encoding that injects information about the order of the sequence, feedforward neural networks that are applied in each layer and layer normalization and residual connections to stabilize and speed up training. Transformers can

be stacked into encoder and decoder blocks. The first is to process input sequences to transform them into a series of context-aware vector representations. Each layer consists of multi-head self-attention, a feedforward layer, residual connections and layer normalization. The decoder generates the output sequence by incorporating previously generated tokens and the encoders output representations. Each layer includes multi-head self-attention, encoder-decoder attention and a feedforward network. the decoder produces the output sequence incrementally. Transformers are trained in parallel which allows for efficient large-scale training. This model scales extremely well and forms the basis of more well-known models such as BERT, GPT, T5, etc. It is used in NLP, computer vision, audio, bioinformatics, etc.

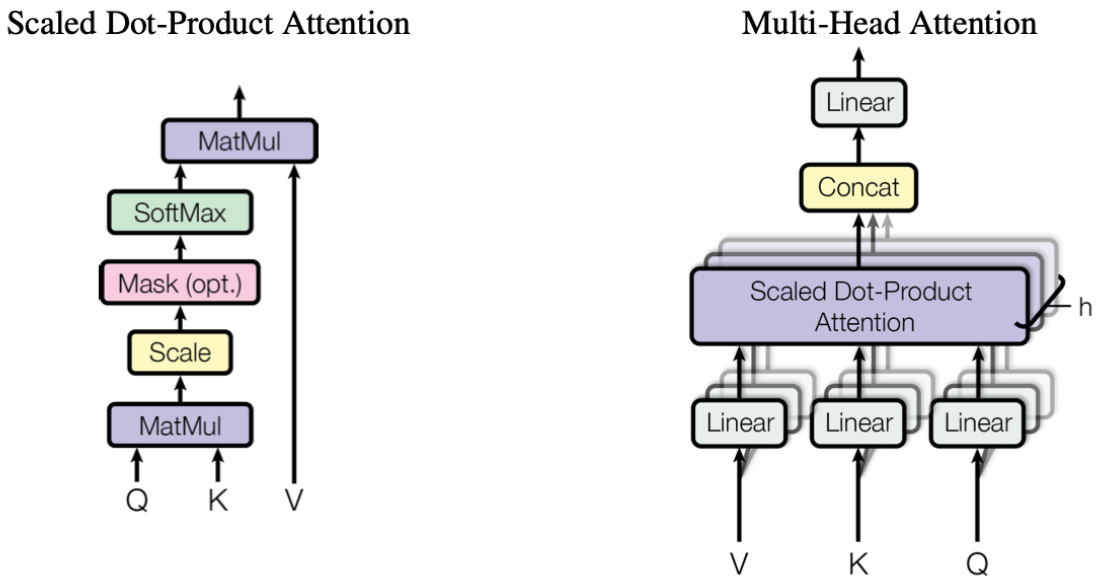

Figure 6. (left) Scaled Dot-Product Attention. (right) Multi-Head Attention consists of several attention layers running in parallel (Vaswani et al., 2017)

Together, these methods provide a robust framework to address the challenges associated with analyzing of biomedical signals. Active learning reduces the need for large amounts of labeled data and improves the model's ability to adapt to each patient. Meanwhile, the neural network models work to find patterns within the data to obtain reliable heart signals for non-invasive monitoring.

Chapter 3. STATE OF THE ART

One of the earliest examples of remote medical devices is continuous glucose monitoring. In the second iteration of that device, the Freestyle Libre 2, developers recognized the need to integrate medical alerts for patients (Miller, 2020). The next step was to identify a classifier capable of notifying both medical professionals and patients when necessary.

Failure detection in ECMO systems has traditionally relied on clinical expertise and predefined monitoring thresholds, which often results in delayed identification of complications. In contrast, recent advances have introduced data-driven approaches that leverage machine learning for predictive analyses. Prior research on remote pulmonary assist systems has demonstrated the effectiveness of anomaly detection techniques in enhancing system monitoring, highlighting the potential of ML-based models to improve failure detection in ECMO systems.

Specific to this project, previous work from this lab has focused on using ML to predict ECMO-related failures by analyzing parameters such as pressure drops and oxygen transfer rates. Despite these advancements, existing models often require extensive labeled datasets, which are costly and time-consuming to obtain. My research seeks to mitigate this limitation by implementing active learning strategies that optimize data labeling efficiency while maintaining high predictive performance. Other related research has demonstrated the application of active learning in various medical domains, such as radiology and electrocardiogram anomaly detection. For instance, active learning has been used in electrocardiogram (ECG) anomaly detection to improve diagnostic accuracy, employing models like LSTM autoencoders for better performance in anomaly detection tasks (Dutta et al., 2021). The adaptability of active learning in medical applications suggests its potential for broader adoption in clinical decision support systems. Studies have also explored hybrid models that combine rule-based methods with machine learning to create robust early-warning systems for ECMO failures. One such model integrates machine learning with rule-

based systems derived from clinical expertise to improve the prediction of medical events and conditions (Vaswani et al., 2017).

Active Learning models include a Machine Learning internal model and a query strategy to properly select samples of the input signal. The machine learning model that was used is a Decision Tree, after a preliminary comparative analysis of standard ML algorithms including: SVM, Gaussian, Decision Trees, Random Forests, Neural Networks, Bayes, QDA and Logistic Regression. For the query strategy there are several methods described in the literature: Stream Random Sampling (Cacciarelli et al., 2023; Lahiri & Tirthapura, 2018; Sibai et al., 2016), period sampler (Sibai et al., 2016), split method (Sibai et al., 2016), Density-based (Bhattacharjee & Mitra, 2021; Mayabadi & Saadatfar, 2022), Cognitive Dual Query Strategy with Random Variable Uncertainty (Liu et al., 2023), Stream Probabilistic AL (Kottke et al., 2015). These methods are described in the section of Methodology and their performance is shown in Results.

In addition to the development of the active learning model for real-time monitoring of the ECMO system, this thesis has included additional algorithms for monitoring vitals of the patient that included respiratory rate and heart condition. For breathing rate estimation, the model will compare a methodology based in peak-to-peak measurements in the time-domain with obtaining the frequency from an FFT. A comparison of these approaches of analysis applied to biomedical signal is presented in (Skotte & Kristiansen, 2014), which concludes that time-domain peak detection is generally a better choice.

There are several ways to stablish the health level of the heat, depending on the type of signals analyzed: Electrocardiogram (ECG or EKG), Seismocardiogram (SCG), or Phonocardiogram (PCG). The correlation between signals that can be obtained from the heart are described and analyzed in (Choudhary et al., 2019), which is the paper in which this part of the research is based on.

Chapter 4. PROJECT DEFINITION

4.1 MOTIVATION

Early failure prediction in ECMO systems is critical for preventing adverse patient outcomes and reducing the burden on healthcare providers. Traditional ML models require large volumes of labeled data to achieve reliable predictions, but in medical applications, acquiring annotated data is challenging due to ethical and logistical constraints. Active learning presents a viable solution by selectively querying the most informative data points, thereby reducing labeling efforts while improving model accuracy and adaptability to different patients and scenarios. Fluctuations in the data may occur because the patient changes posture (lying down, walking, or doing light exercise) or because of changes in environmental conditions. The model must not detect these different scenarios as anomalies, which is why using active learning (AL), the model can learn as it goes and permanently adapt to each patient and situation. Additionally, each patient will have a different base line and 'normal' state, which is another reason the model must be able to adapt after a short autonomous learning period.

The potential clinical impact of this research extends beyond ECMO systems, contributing to broader efforts in intelligent medical device monitoring. By enhancing real-time anomaly detection capabilities, this project aims to support timely clinical interventions, optimize ECMO resource utilization, and improve patient safety.

Beyond improving ECMO monitoring, this work has implications for future biomedical engineering applications. The ability to use machine learning to detect mechanical failures in real-time can be extended to other life-support systems, such as ventilators and dialysis machines. If successful, this research could lay the groundwork for more advanced AI-assisted monitoring systems in intensive care units.

4.2 OBJECTIVES

The overall goal of this project is to monitor a patient using a portable ECMO machine to alert them and their doctors with enough time to fix any problems that may arise. Originally, the scope included only sensors attached to the ECMO machine (pressure and flow), but throughout the project, other vitals were included into the monitoring system to make it more comprehensive (breathing rate and heart signal). All the objectives are described in more detail in the following subsections.

4.2.1 INITIAL PROJECT OBJECTIVES

The primary objective of this project is to develop a machine learning-based algorithm capable of predicting ECMO failure events, with a particular focus on clot formation. The model was initially created in a laboratory setting using water pipes to simulate blood flow, and then the model was validated in a clinical experiment installing an ECMO device in a life sheep (Figure 7).

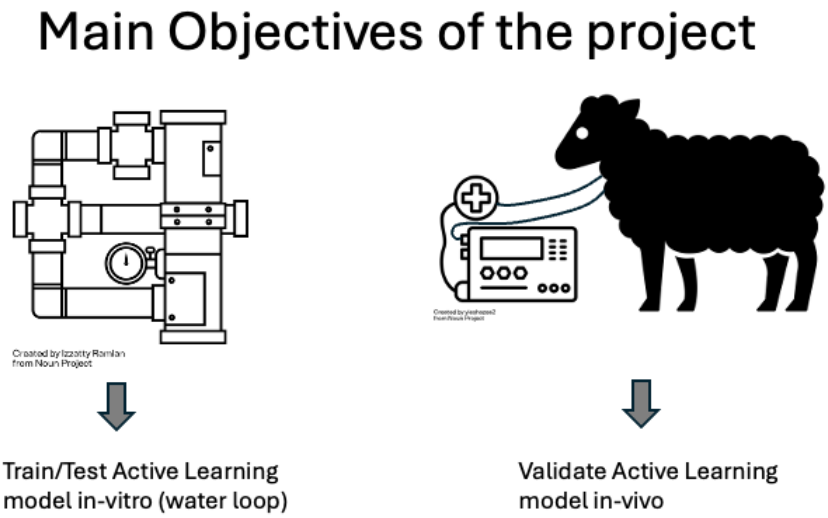

Figure 7. Main objectives involve in-vitro and in-vivo testing

In order to enhance model generalizability, active learning techniques were implemented to minimize the dependency on large, labeled datasets and to automatically adapt to the changing distributions between in-vitro and in-vivo data, and between different patients. Furthermore, the project seeks to validate the model's performance through in-vitro and in-vivo testing, ensuring its applicability in clinical environments. To develop a reliable monitoring system, controlled experiments will be used to generate training data that accurately represents real-world failure scenarios. With this data, the model can be trained using controlled failure cases that simulate what would happen in-vivo. Finally, the model was tested on in-vivo experimental data, installing an ECMO device to a sheep in a controlled environment.

A future goal will be to investigate the potential for integrating this technology into existing hospital infrastructure to facilitate seamless adoption into clinical workflows, considering technical and ethical issues.

4.2.2 ADDITIONAL OBJECTIVES

Throughout this project, two additional vital signs were recorded to help in the overall monitoring of the patient. Having a more comprehensive view of the patients will help in making more accurate predictions and it gives doctors more information. From the experiments with sheep, a breathing rate estimator was developed. Then a machine learning model was developed to estimate the seismocardiogram (SCG) heart signal, a signal that doctor use to evaluate the health of the heart, using the sound recorded with a smartphone and headphones. To develop this model, the recordings were obtained in a controlled setting using my own data (human data instead of sheep data). Additional objectives are depicted on Figure 8.

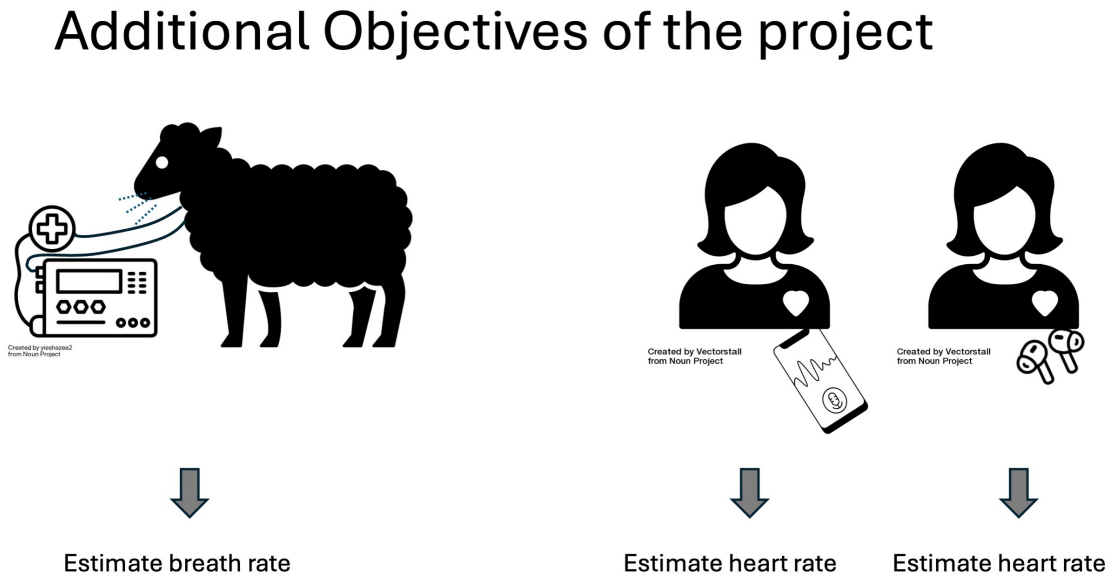

Figure 8 Additional objectives involve estimating vital signals

The first vital sign added was **breathing rate**. When the in-vivo sheep studies started, the data showed a new pattern that did not exist in the previous in-vitro (see Figure 9). This pattern can be seen in the pressure sensor data after zooming in, and it is very periodic. The first thoughts were that it could be related to the heart rate, as it would show a higher increase of pressure when the blood is being pumped out of the heart. However, based on the positions of the sensors and the rate of repetition, this idea was discarded. After talking to three of the doctors and biomedical staff in charge of the experiments, these repetitions were said to be the sheep's breathing rate. Calculating breathing rate of the patient can be informative of whether the patient is resting, doing exercise, having trouble breathing, or any other similar situations. The goal of monitoring breathing rate through the inlet pressure sensor was then added due to its informative value to the complete system.

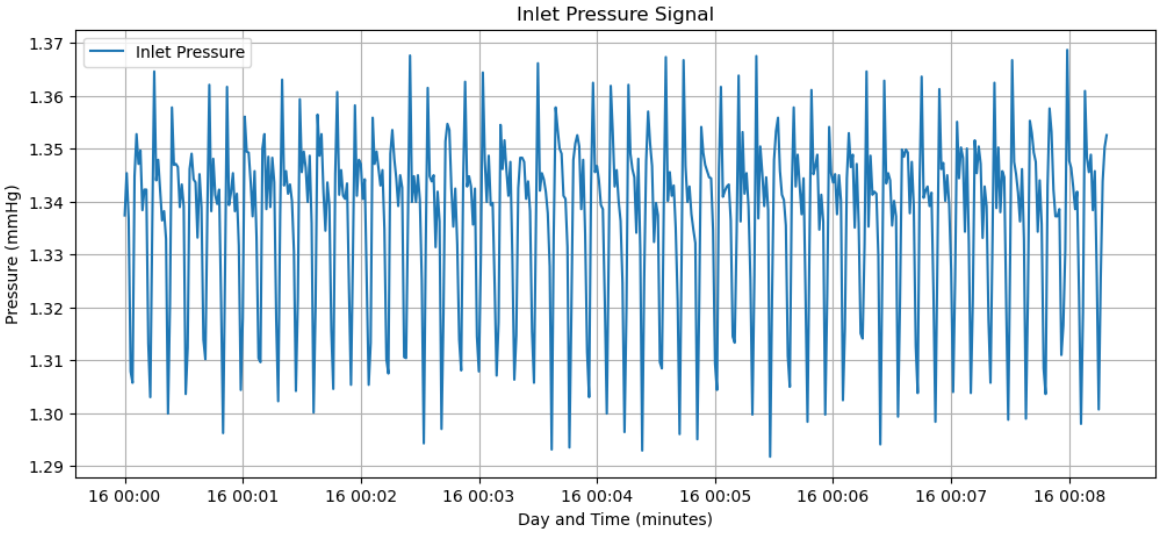

Figure 9. Pressure sensor from in-vivo sheep data where a repetitive pattern is visible

The second additional objective was recording the patient's **heart signal** for the same main reasons mentioned previously, to obtain a more complete monitoring structure for the patient. The most informative heart signals are electrocardiograms, echocardiograms and seismocardiograms, all of which require specialized equipment that require going to the hospital. This last objective concentrates on obtaining a reliable seismocardiogram signal from the patient in a non-invasive way. This would allow for a dependable and easy way of monitoring the patient's heart without the need for them to go to the hospital. To accomplish the goal of developing a non-invasive and easy to use procedure, the signal is recorded with in-ear headphones as a phonocardiogram (PCG) and then mapping that signal to a seismocardiogram (SCG). PCG records sounds (similar to using a stethoscope), while SCG measures the hearts movement. By mapping PCG to SCG we can effectively obtain one of the most informative heart signals thought recording sound with in-ear headphones.

4.3 METHODOLOGY

Throughout this project, several models have been developed with the main objective of building a system which is able to monitor, detect and predict failure events in a portable ECMO machine. These models have been developed with active learning techniques, which allow for a reduced labeled dataset and can adapt to different distributions in real time. The first model was trained with in-vitro data that came from a water-loop circuit simulating blood flow in the machine. Several sets of experiments were conducted to simulate sudden (Figure 10) and slow (Figure 11) pressure decreases. A slow pressure decrease can indicate a problem within the system like a blood clot over the oxygenator, while a sudden pressure decrease might mean there's kinking in the tubing or other malfunction.

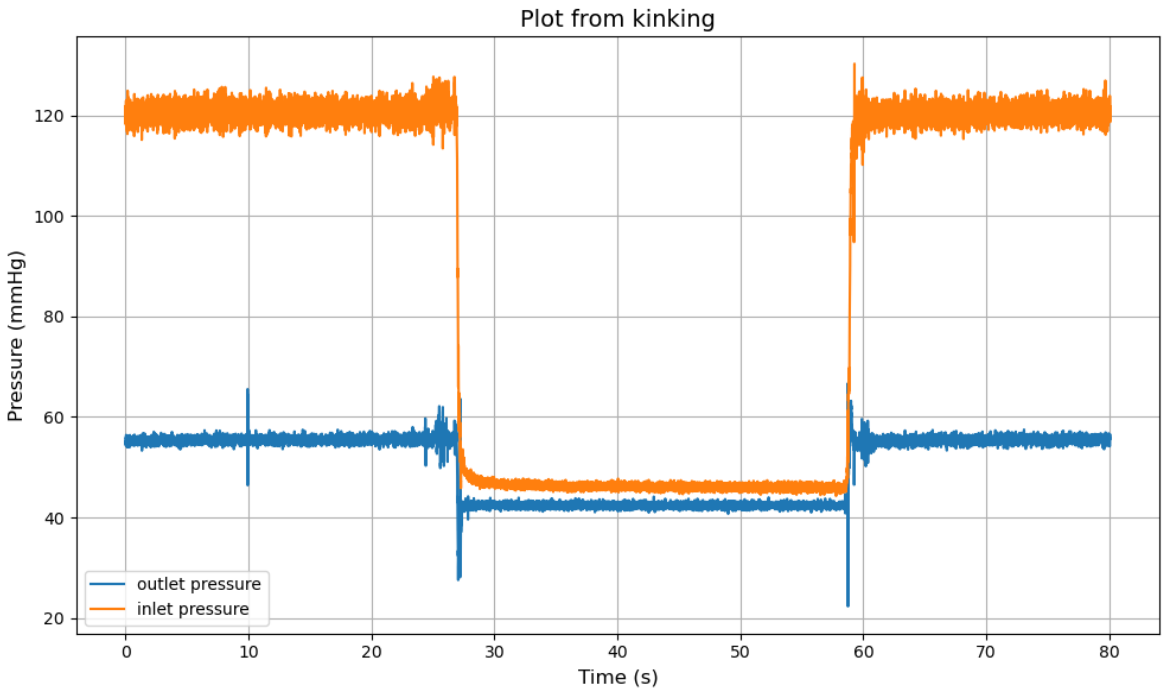

Figure 10. Sudden pressure drop simulated in the water loop circuit

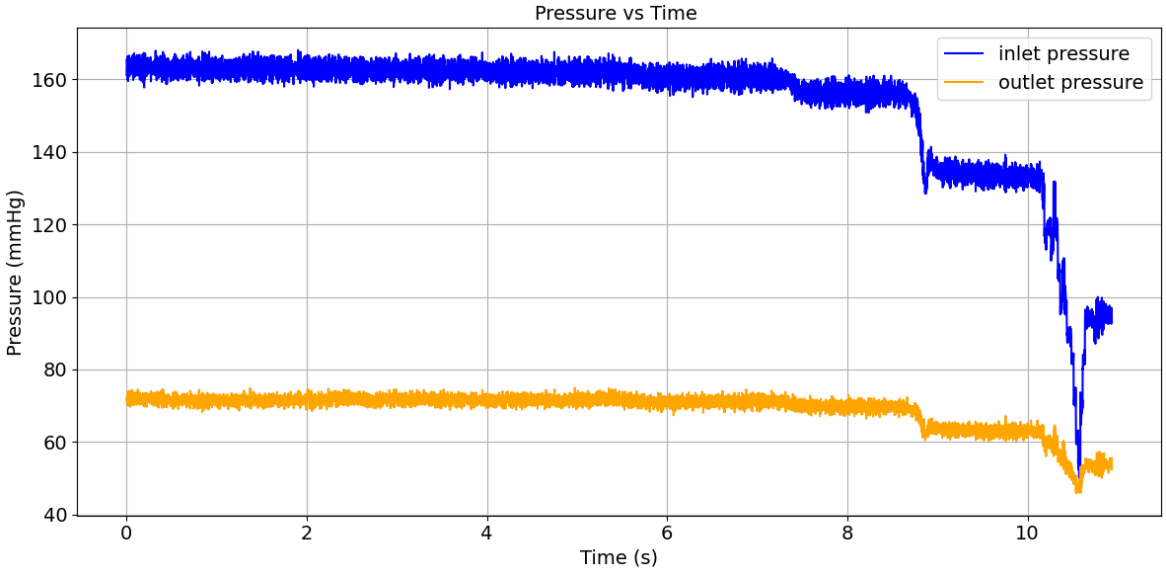

Figure 11. Slow pressure decrease simulated in the water loop circuit

The failures were labeled when the pressure decrease was considered to be too abrupt, or low enough to be of concern. Afterwards a model was trained to detect anomalous pressures and querying the most informative data points.

In a second phase, sheep experiments were conducted, and the in-vivo data was obtained. Two datasets came from the hospital and served different purposes. The first one shows a series of pressure and flow drops over time. These failure events were labeled as such, marked in red in Figure 12, and were used to two models. The first is a pool-based active learning model to compare eleven basic machine learning models with the same query strategy. After this comparison, decision trees were chosen as a main model. The second model that was developed was a stream based active learning model to compare six query strategies in order to optimize anomaly detection.

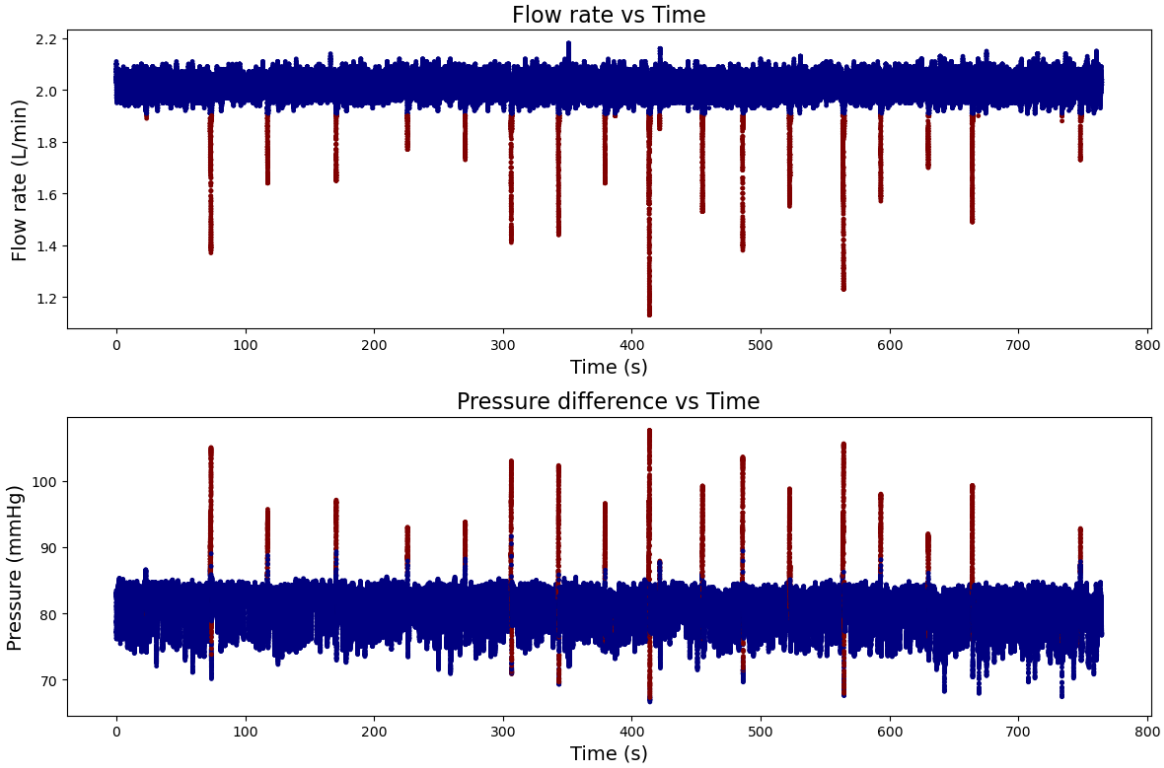

Figure 12. Graph comparison of data from the flow sensor and the pressure sensor where the suction events are aligned

The second set of data contains only pressure data as the rest of the sensors malfunctioned during the experiments. The sheep survived the experiment without any complications; therefore, the data has no notable anomalies. During the analysis of this data, I saw that there was a repetitive pattern that showed the sheep's respiration. This motivated me to develop a new model that calculates and monitors breathing rate in real time, allowing for an increased monitoring of the patient.

Lastly, a model was developed to obtain heart signals (SCG) through audio recorded from in-ear headphones. The signal obtained with headphones is informative as it contains phonocardiogram signal (PCG) and can be mapped with the developed model to a more informative SCG signal that the doctors are used to analyze. To train and validate the model, synchronized PCG and SCG data was recorded on a phone placed on my chest. The data was preprocessed aligned and cropped into individual heart beats. Three neural networks were

trained: a convolutional neural network, a recurrent neural network and a transformer to solve the regression problem of mapping PCG into SCG. This allows for a non-invasive and practical way to monitor the patient's heart signal and completing a well-rounded monitoring system without the need of specialized medical instrumentation.

4.4 PLANIFICATION AND FINANCIAL ESTIMATE

The research follows a structured methodology comprising multiple experimental tasks described below:

Data Acquisition and Signal Analysis. Initially, suction event detection is conducted by analyzing flow and pressure sensor signals to determine whether active learning can enhance anomaly detection. This data is obtained from in-vitro testing on a water circuit created to simulate the portable ECMO machine. The work described in this thesis does not include building the physical simulator but starts with the analysis of sensor signals obtained.

Development of Algorithms and Validation. A subsequent phase involves implementing and validating algorithms in realistic settings. This includes in-vitro testing with blood circuits. Experimental setups include controlled failure scenarios using simulated blood circuits to train and refine ML models. Once proven that AL provides an advantage versus other ML strategies, the model must be validated on failure cases. The integration of active learning is a key component of the methodology. By utilizing uncertainty sampling and other query strategies, the model iteratively selects the most informative data points, thereby reducing annotation requirements while maintaining robust predictive capabilities.

In-vivo Testing The final step is to test the model previously created by in-vivo studies conducted on sheep. The model's performance is assessed based on its ability to distinguish between normal and failure conditions with high sensitivity and specificity.

Exploring Future Work. If the project finishes ahead of schedule, the research will explore other technologies such as reinforcement learning as a complementary method to active

learning, potentially improving the adaptability of the model. Reinforcement learning could allow the system to dynamically adjust failure prediction thresholds based on real-time data.

Writing Thesis Document. It is planned to write the thesis document during the development of the project. Possible scientific publications will depend on the quality of the results obtained.

Task	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
Data Acquisition and Signal Analysis									
Development of Algorithms and Validation									
In-vivo Testing									
Exploring Future Work									
Writing Thesis Document									

Figure 13. Chronogram of activities

Financial estimate. The main cost of the project would consist of the hours worked, as the materials and experiments all came from the Carnegie Mellon labs.

The number of hours dedicated to the research presented in this thesis was 400 hours. This time includes the writing of the thesis document and slides for the dissertation, as well as follow-up documents and presentations during the development of the project. However, it doesn't include the time devoted to writing the papers submitted to conferences and technical journals.

The volume of data, number of signal and complexity of the models did not surpass the capability of current laptop computers with CPU and GPU capabilities. Therefore, there was no need to use cloud server or High-Performance Computer (HPC) to develop this project.

Building the ECMO prototypes, preparing the water simulator in the lab, and running the in-vivo experiments with real sheep involve important costs in the big project of developing autonomous ECMO devices. However, the research presented in this thesis is focused on the analysis of the signals, and the development and testing of the models.

For the additional objectives related with heart signal analysis, the phonocardiogram (PCG) and the seismocardiogram (SCG) were collected using a smartphone and in-ear headphone, reading accelerometer and microphone signals. Hence, no additional expenses were required.

Chapter 5. DEVELOPED SYSTEM/MODEL

Throughout this project, several models have been developed to create a whole system with the goal of controlling, detecting and predicting failure cases in a portable ECMO machine. All models were developed using experimental data. The first model was developed using in-vitro data from the lab where a water loop was built. The model showed monitoring capabilities on pressure data, and 18 specific events were detected using the active learning approach. Next, data was obtained from in-vivo sheep experiments, and this model was modified to work with two input signals: pressure and flow. From the sheep data the breathing pattern could be observed, so a third model was developed to detect breathing rate and keep an eye on the patient. Finally, to further monitor the patient, a model to analyze heart rate with a smartphone and headphones was also developed. The data for this final model was recorded on myself, using a personal smartphone placed on the chest near the heart. All models developed in this research thesis are summarized in Figure 14, which also indicates the source of the data used for each model: water loop, sheep or human.

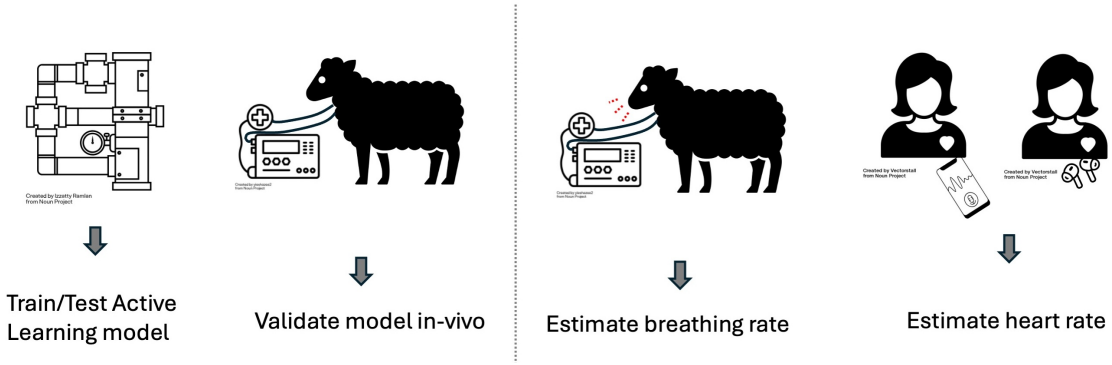

Figure 14. Summary of all models developed. Initial objectives (left part) and Additional objectives (right part)

5.1 ACTIVE LEARNING MODEL ON WATER LOOP DATA

Passive learning, also known as supervised learning, relies on a static labeled dataset to train and test a machine learning model, in a process called supervised learning. However, active learning, which is rooted in information theory, can overcome the limitation of labelling data and also provide de capabilities of self-adaptation to a new patient or new working environment. Active learning identifies the most informative data points and requests that an oracle label them, thereby improving the model's performance with minimal labeled data (Settles, 2009).

The first model was developed on in-vitro data, simulated though experiments done in a water loop installation described below (Figure 17). The goal of these experiments was to simulate what real case failure events might look like in the in-vivo experiments. Two main anomalies were tested, tube kinking and slow pressure decrease. The first one occurs when the tubing is bent or caught somewhere that cuts blood flow within the tube (Figure 15). This can cause an abrupt pressure increase or decrease depending on the relative position of the kinking and the pressure sensor. The second type of anomaly occurs when a blood clot is forming which makes the pressure in some areas of the circuit slowly increase and in others to slowly decrease (Figure 16). This sustained change is not possible to detect with a simple threshold as there are other variables that can lead to changes in pressure, such as if the

patient is moving around and exerting themselves more or laying down sleeping. Additionally, each patient might have a different baseline which affects the levels at which the pressure drop becomes critical. Active learning allows for a progressive learning and adaptation to each situation and patient, querying the correct data samples to know if the pressure change can still be considered within a regular range.

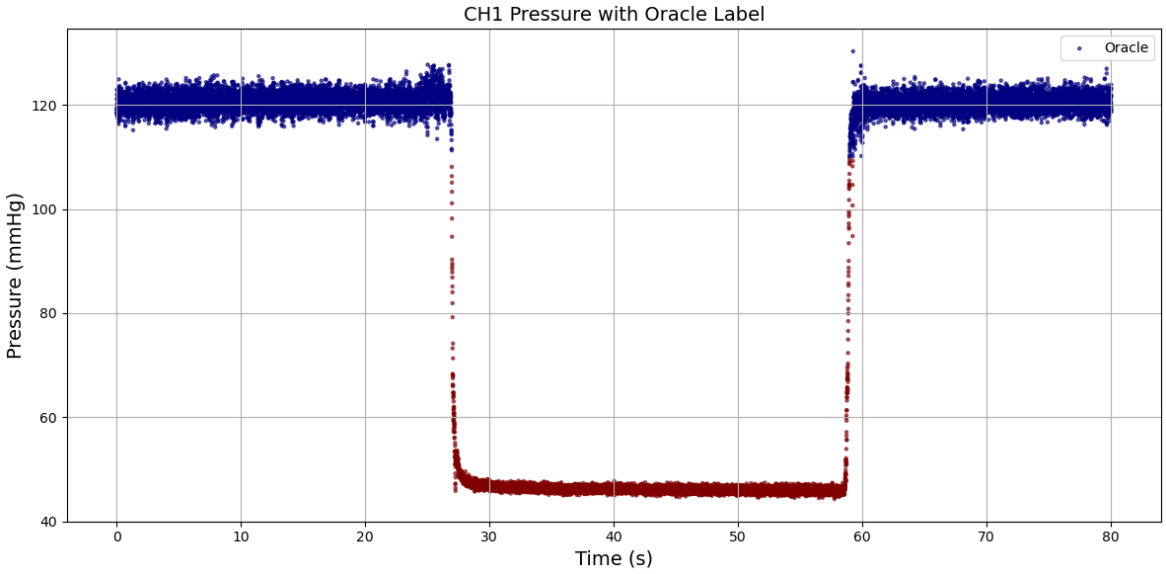

Figure 15. Instant pressure drop, or kinking where red indicates anomaly and blue is regular data

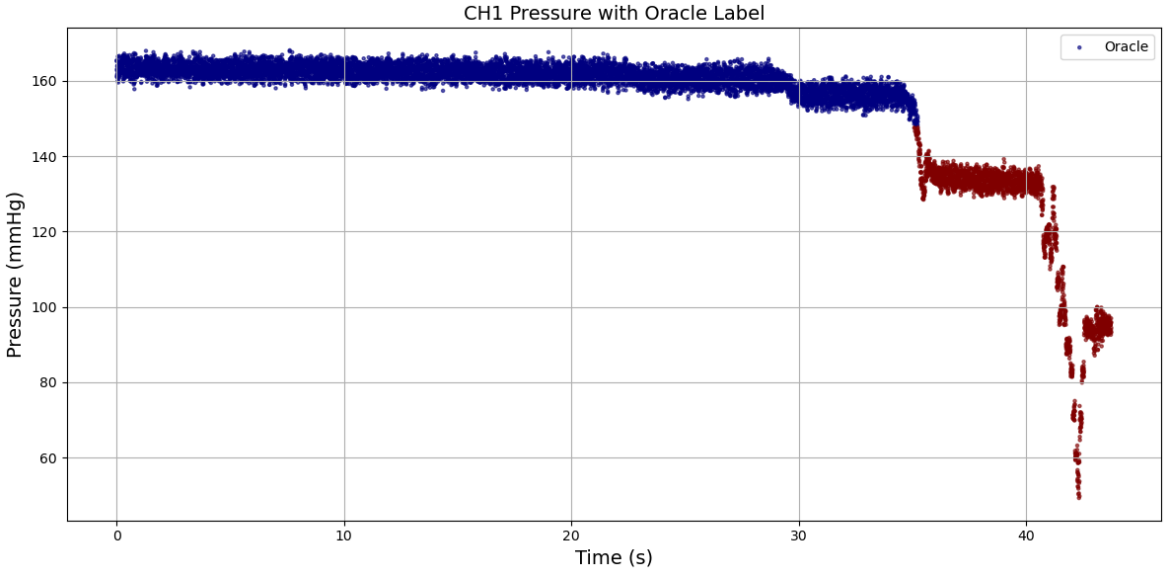

Figure 16. Slow pressure drop over time where red indicates anomaly and blue is regular data

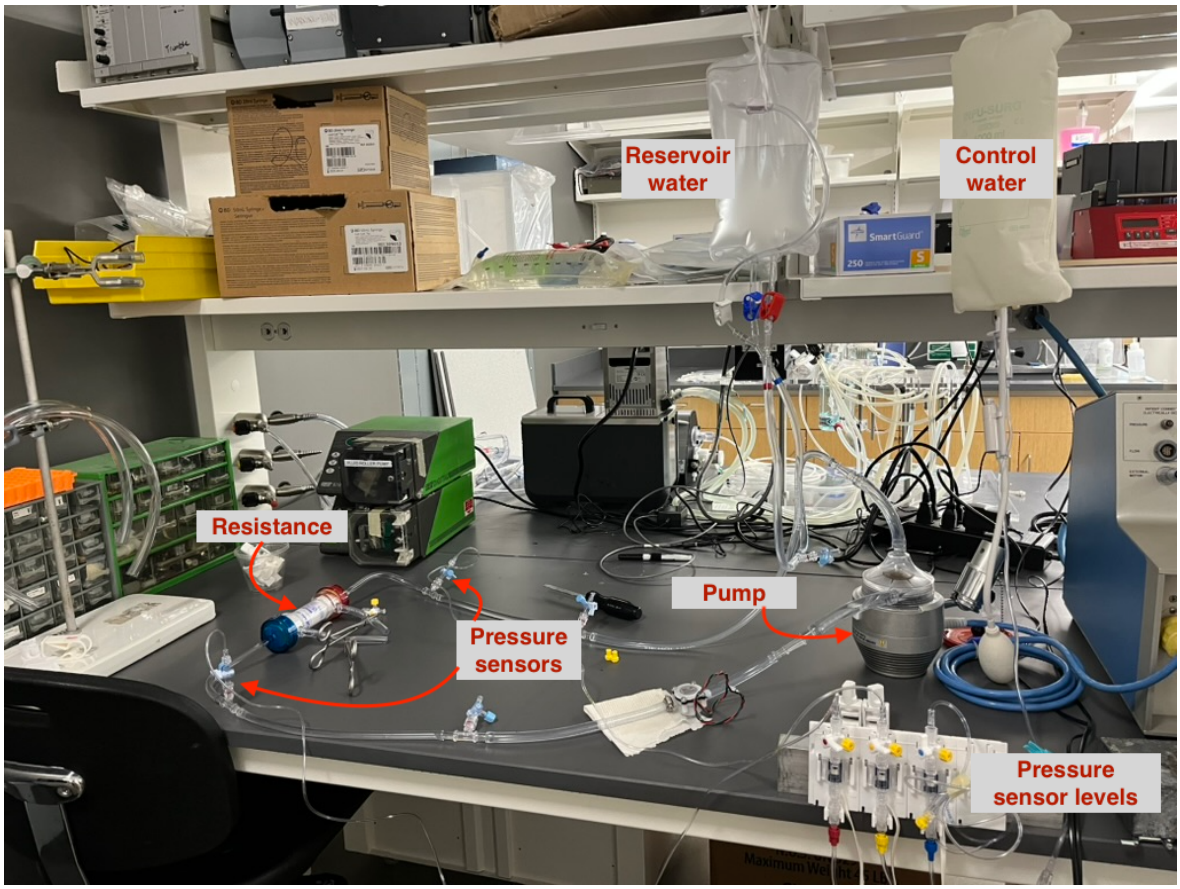

Figure 17. Water Loop experiment circuit

Figure 17 shows the water loop setup that was built in the lab. The pressure sensors can be seen marked on the left on either side of a resistance simulating the oxygenator. On the top right is a bag representing the body with reservoir water. The bag next to it marked 'control water' is used to maintain the correct pressure levels for the pressure sensors its connected to (bottom right). The last element is the pump located on the right-hand side.

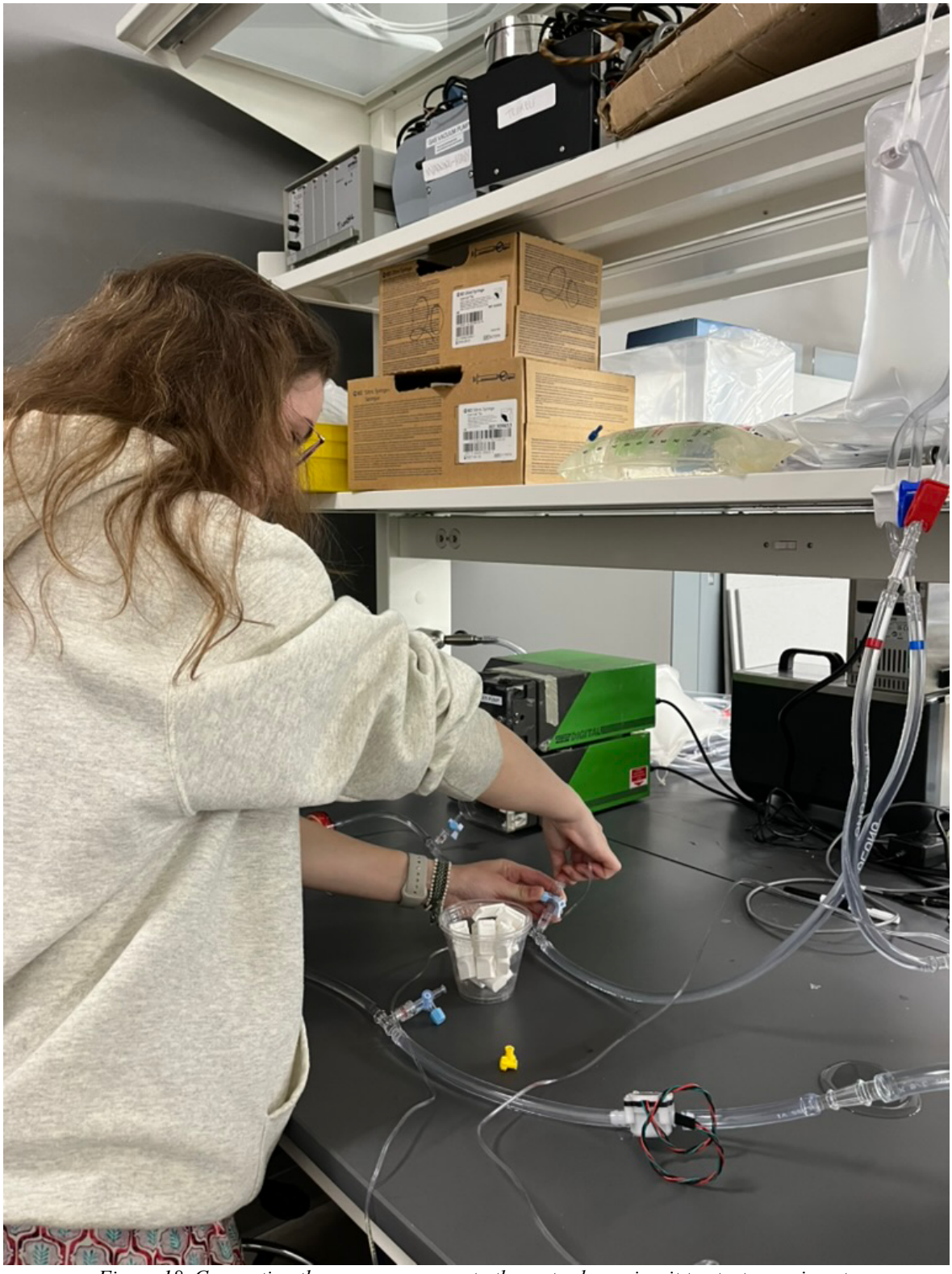

Figure 18. Connecting the pressure sensor to the water loop circuit to start experiments

The circuit was made with several ports to which attach the pressure sensors. Figure 18 shows me attaching one of the pressure sensors to the circuit to calibrate. After measuring the circuits pressure in different places, they were finally installed on either side of the resistance for all of the subsequent experiments. The slow pressure drops were simulated by slowly cutting flow from the tubes. This was done with the claps shown in Figure 19 that compressed the tube as it was turned. This was done progressively until the flow in the tube was completely cut off.

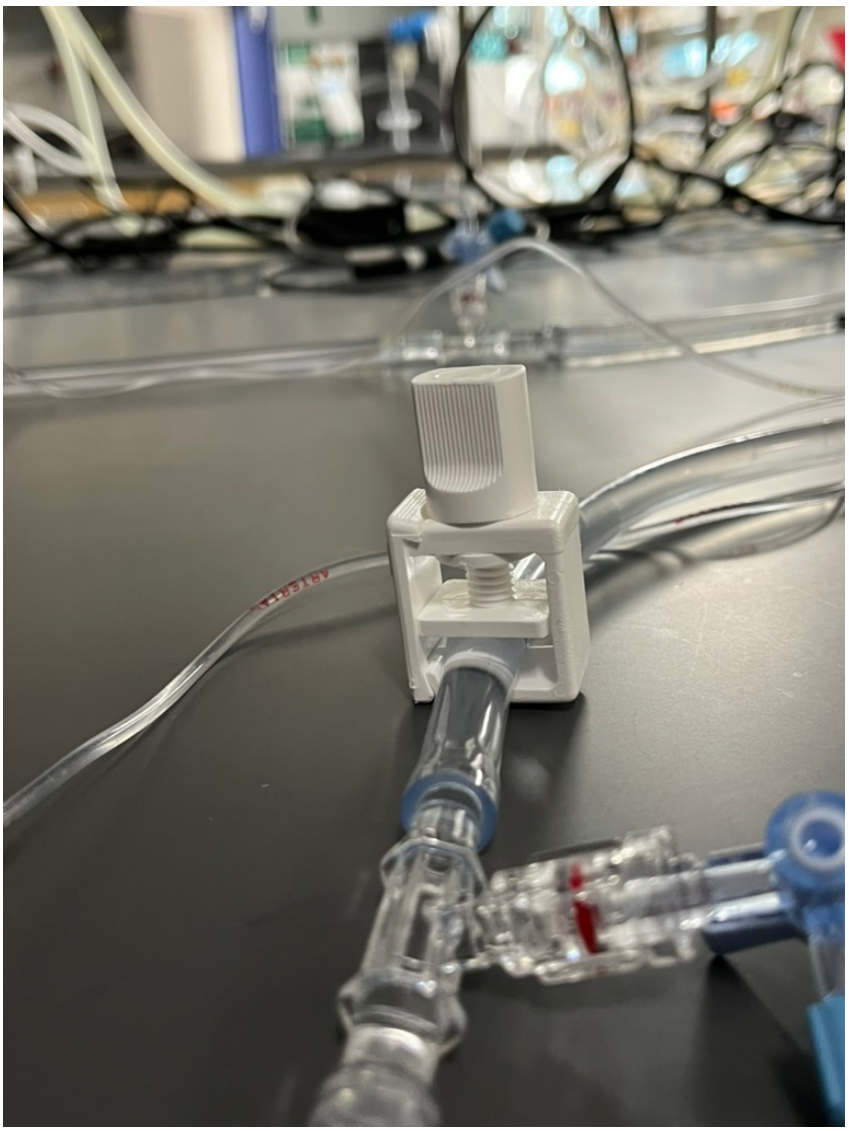

Figure 19. Clamp on tubing of water circuit

To train the best model possible, six different query strategies were implemented and evaluated. Query strategies are algorithms that select the most interesting data points of the signal to learn from. To ensure a fair comparison, all strategies had access to the same time-windows size from which they selected the training points they considered most informative and they all used the same decision tree algorithm. The strategies were then evaluated and compared based on accuracy and acquisition count—the number of samples queried throughout the process. The six query strategies are described below.

Stream Random Sampling (Cacciarelli et al., 2023; Lahiri & Tirthapura, 2018; Sibai et al., 2016) selects random samples from the input stream. This method is simple and easy to implement; however, it may not contain samples from all possible classes. Therefore, manual pre-selection of at least one point from each class or several iterations may be required to ensure that all classes are represented.

The **period sampler** (Sibai et al., 2016) chooses one sample for every n samples. This method is simple to implement and very consistent. However, it is inefficient and inflexible, meaning it might not adapt to changes in the data, and it may miss periodic information because it may never be sampled. A previous analysis of the data is necessary to select an appropriate value for the parameter n.

The **split** method (Sibai et al., 2016) divides the stream into fixed-size segments. Some instances are then sampled using other methods (such as random or periodic sampling) for each segment. It's simple, systematic, and balanced, but inefficient and rigid.

Density-based (Bhattacharjee & Mitra, 2021; Mayabadi & Saadatfar, 2022). This method is based on the idea that instances in denser regions of the feature space are more informative. It avoids outliers and is efficient and adaptive; however, it has a high computational cost and may overlook rare or important data.

Cognitive Dual Query Strategy with Random Variable Uncertainty (Liu et al., 2023). This method selects data points based on random variable samples and uncertainty sampling. It selects the data points that the model is most uncertain about, as estimated by entropy,

margin, or variance prediction. The cognitive aspect of this method helps guide the strategy and simulates human decision-making or information acquisition. It is efficient and balanced, but complex and computationally expensive.

Stream Probabilistic AL (Kottke et al., 2015). In this method, a probability of query is assigned to each instance based on uncertainty (the model's confidence in its prediction), informativeness (the model's potential to improve), and diversity (the degree to which an instance represents a new or underrepresented region). Instances are randomly queried, and the probability value changes with the model and new data. This method is flexible, adaptive, and efficiently labels data. However, it has a high computational cost and risks missing key instances.

The comparison all these methods in terms of accuracy, along with the ability of the Active Learning model to detect the induced anomalies will be shown later in the Results section.

5.2 ACTIVE LEARNING MODEL ON SHEEP DATA

Following the training of an active learning model on in-vitro data generated in the laboratory, the model was evaluated using in-vivo data from clinical experiments on sheep. These experiments spanned 14 days, during which a portable ECMO machine was attached to the sheep, and data was recorded on four sensors: two for flow measurement and two for pressure. It is imperative to clarify that the sheep did not succumb to death or experience severe discomfort or distress during the experiments. They were able to sit, sleep, or stand at their convenience, and their vital signs remained stable throughout the duration of the experiments.

Three experiments were conducted on sheep. The initial experiment was utilized to assess the performance in-vivo of the active learning model. The data from the subsequent two experiments was used to of identifying any noteworthy events. In the second experiment, we had access to both pressure and flow data for the entire duration of the experiment. The data

collected was mostly noisy, due to ongoing refinement of the portable ECMO machine, which resulted in some sensors not recording data accurately during the experiment. However, it was possible to collect flow and pressure data accurately during several hours. The final experiment yielded two days of viable data. The data collected in all three experiments was very useful to demonstrate the capability of the model to detect anomalies, to analyze changes in position of the sheep and to develop the breathing rate model.

5.2.1 FLOW MODEL

During the first experiment, while the system was being set up, some problems arose in the tubing and several adjustments were made in the moment. This resulted in data containing 17 suction events during 12.7 minutes (Figure 20) that were very useful as a simulation of anomalies without putting the sheep in any risk. The total number of samples collected was 191,130. Decreases in flow due to the suction events are clearly visible as the speed decreases from a stable 2 l/min and these flow drops are considered anomalies. These events are then marked as failure cases (red) for training and testing the model and can be seen in Figure 20.

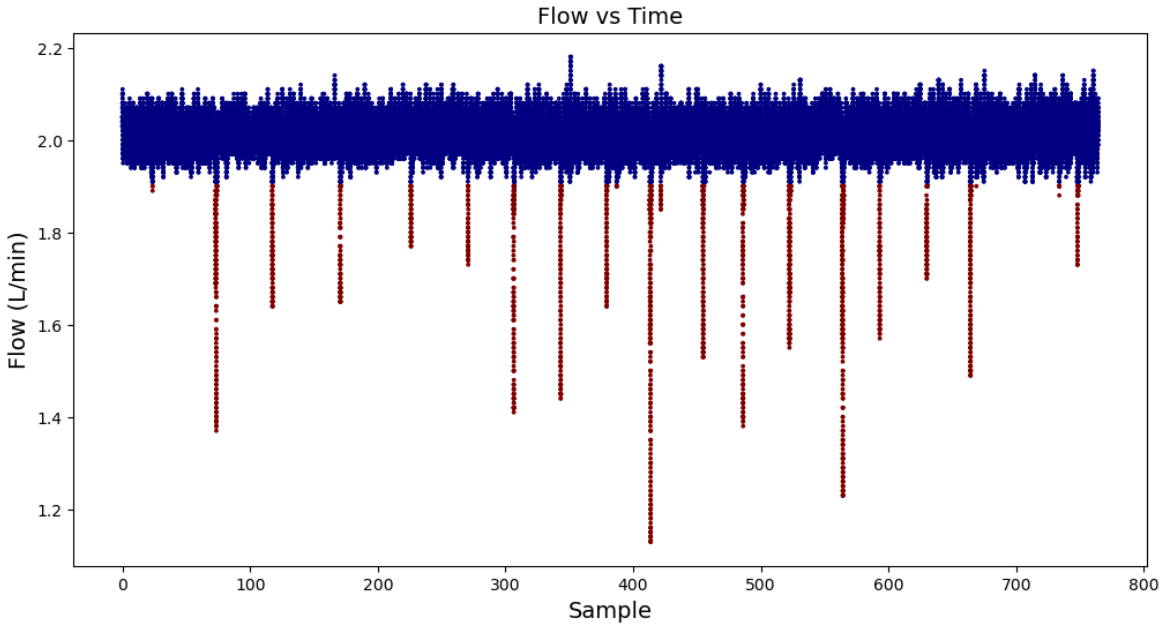

Figure 20. Blood flow during 12.5 minutes. Suction events in flow marked in red

Before applying real-time active learning techniques, a comparative study of eleven traditional machine learning algorithms was conducted. Using the Pool-Based approach (Cacciarelli et al., 2023), in which the query strategy has access to all the training data point to then select a subset of points to train a model, then the model was evaluated for accuracy using the held out labels.

For the comparative analysis of machine learning models, subsets of different sizes were used, comprising 2, 10 and 20 samples were randomly selected from the entire dataset. Since real implementations of the active learning approach require a short training time for fast adaptation to variable conditions, the number of samples used to train the model must be low (Cacciarelli et al., 2023). The algorithms compared were well-known machine learning methods: SVM, Gaussian, Decision Trees, Random Forests, Neural Networks, Bayes, QDA and Logistic Regression, out of which Decision Trees were picked for the next step.

The next step was to compare the 6 query strategies from the previous model. The same procedure was followed as in the case of in-vitro data, where all the queries were evaluated

in the same conditions. The Results section contains all the details of the outcomes of these analyses.

5.2.2 FLOW AND PRESSURE MODEL

In order to increase performance in this model, the input data was modified to contain both flow data and the pressure difference. Initially the models were using just one input valiable, but since data was collected for flow and pressure during the experiments with the sheep, a new model was developed to include both inputs. In this second set of data, the pressure suddenly increases during suction events that produce a drop in flow, as can be seen in Figure 21. The flow data was manually labeled, and those same labels were applied to the pressure difference data. Marked in red show all data points considered anomalies that the model must learn to detect.

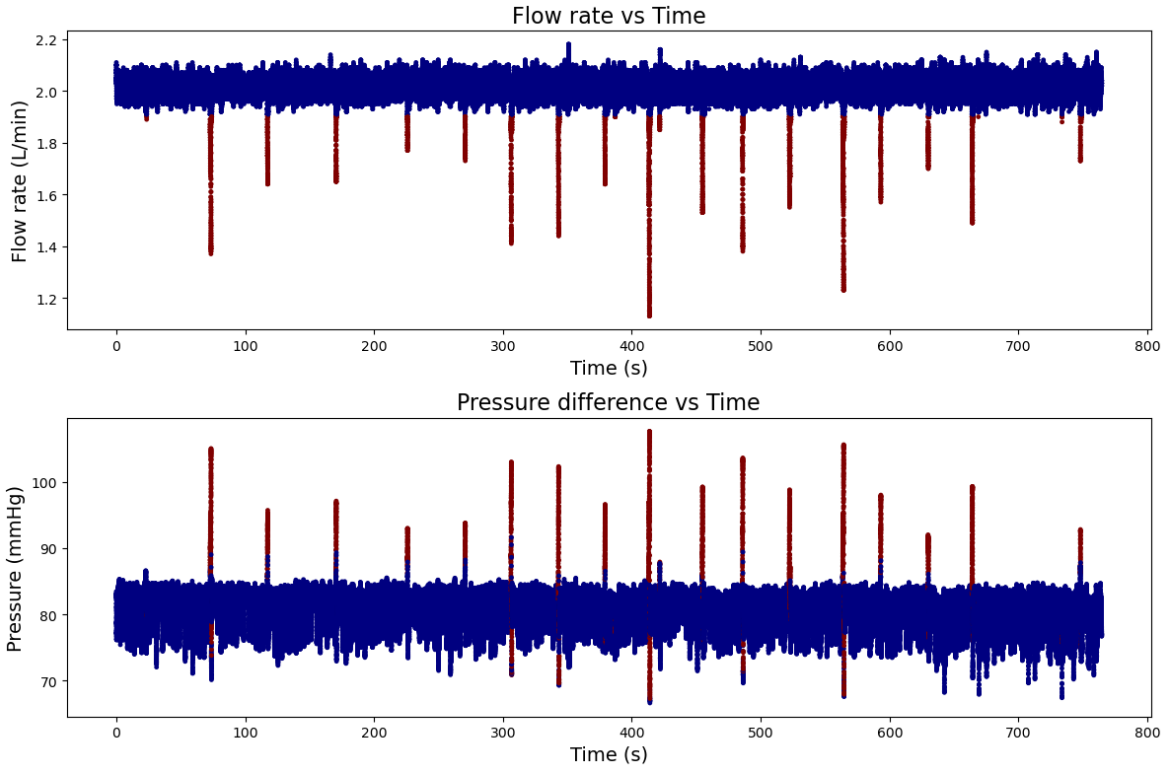

Figure 21. Suction events in both flow data (top) and pressure difference data (bottom)

The remaining data from the second and third experiments was combed to find any interesting patterns and filtered to remove the noise. The most interesting observations were seeing the changes in pressure when the sheep moved, as well as seeing its breathing rate. It's interesting to see that movements produce small pressure changes that are reflected in the data. As shown in Figure 22, there are downward events at around 8:00, 8:20, and 9:00. This is not a very distinctive pattern (the flow drop is just 4%) that should not trigger an anomaly in the monitoring system. When mentioning this to the team in charge of the sensor set up on the sheep, they mentioned that the sheep could only stand up or lay down and that it changed position a couple of times per hour. Each movement is shown in the flow data as a peak that after a couple of minutes returns to its previous value. This is where active learning becomes useful as it still be able to adapt the model and not sound alerts at small and normal changes due to variations in environmental conditions as opposed to pressure/flow variations due to upcoming malfunctions.

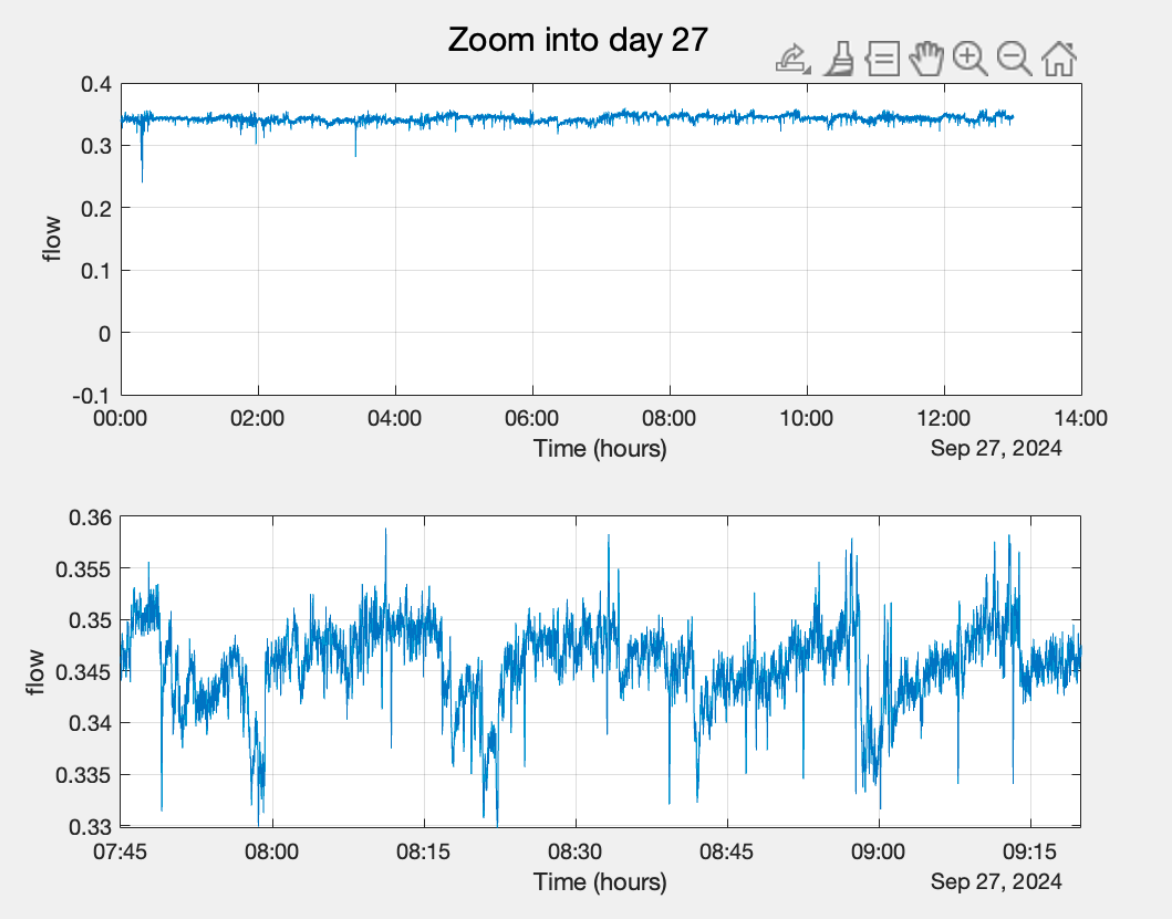

Figure 22. Pattern seen on the last day of in-vivo testing around every 20 minutes

The second observation on the pressure signal was a much quicker pattern, repeating about 6 times per minute. When talking to the biomedical team, they suggested this might be the sheep's breathing rate. This pattern can be clearly seen in Figure 23. Further analysis on this characteristic of the signal yields to a new model described below.

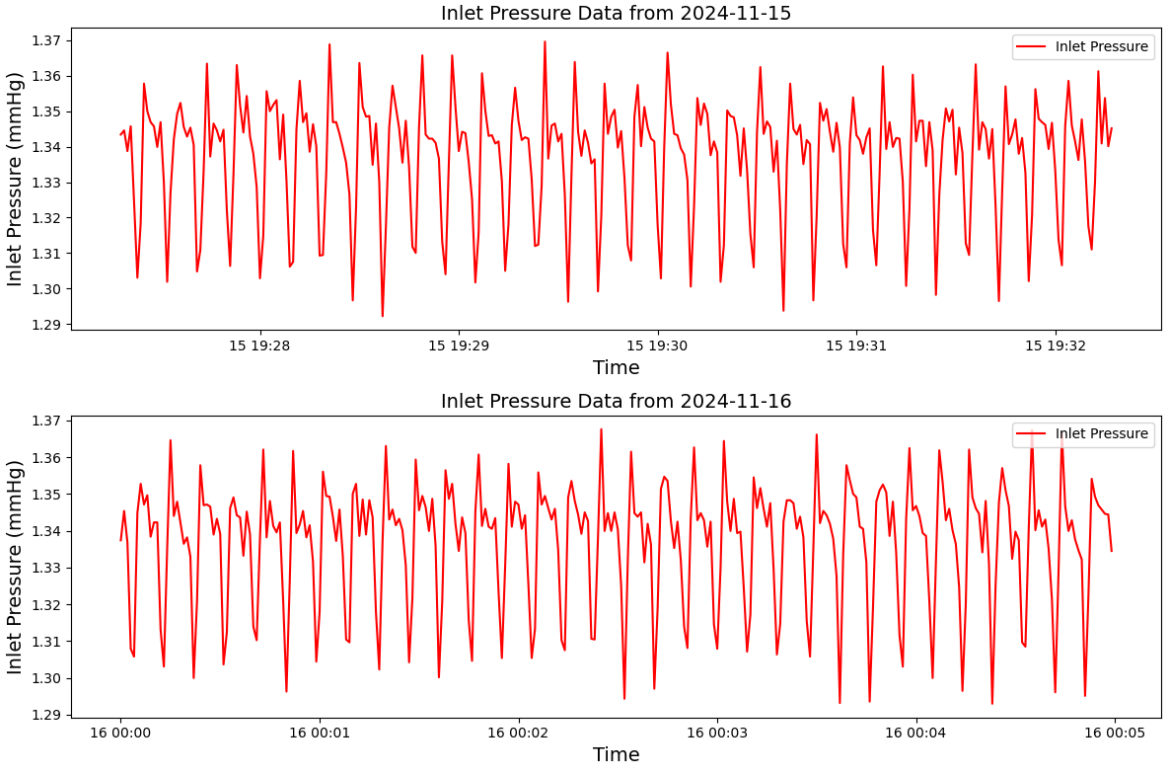

Figure 23. Zoomed in version of inlet pressure data from two separate days in in-vivo sheep trials

5.3 BREATHING DETECTION MODEL

Although it was not previously planned for this project, an additional model was developed to process and detect the sheep's breathing rate in real time. When analyzing the complete in-vivo data from the sheep experiments, a distinct patter could be seen in the pressure data (Figure 23). This pattern has a very slow periodicity compared to a flow pump, which meant that it had to be associated to a repetitive process in the sheep's body. Heart rate was discarded because it would have had to be roughly ten times faster frequency, which left breathing as the only other viable option, as it was later confirmed by the medical team. Detecting breathing would allow for a more well-rounded monitoring of the patient. Increased breathing rates may indicate the patient is doing a more demanding activity. This, coupled with changes in the rest of the data may mean that the patient is changing activity instead of the machine failing.

To detect repetitive patterns with peaks, there are two approaches: peak detection and applying a Fourier transform to analyze frequency. Both methods were compared and evaluated with different sized windows in the Results section.

5.4 HEART SIGNAL DETECTION MODEL

The final step, a more complete and non-invasive monitoring system was developed to obtain heat rate and heart signal. There are different signals associated with heart health, depending on how they are obtained. The most common, important and most informative is the Electrocardiogram (ECG) which records electrical activity from the heart. It is considered the "gold standard" for cardiac diagnosis. Seismocardiography (SCG) is a non-invasive heart measurement that records the vibrations of the chest wall induced by the mechanical activity of the heart. These vibrations can be captured by accelerometers placed on a patient's chest. SCG gives information about the mechanical aspect of the heart, its valves, contractions, etc. Doctors check for valve openings and closing, the amplitude and shape of peaks and beat consistency. Our experiments show that this signal can be captured with a smartphone's accelerometer by just placing it on the chest of the patient, with the bottom of the phone towards the chest near the heart as seen in Figure 24. This measuring technique does not need specialized equipment like an ECG, but it can be done by any individual with a smartphone and without any special training. However, this requires the patient to be completely still during the complete recording, which is not an easy feat, nor is it comfortable enough to do frequently. For our study, this data was collected on myself as the sole subject, using a personal smartphone. The recordings were performed voluntarily and with informed consent, and no identifying or sensitive information was stored. All data were handled securely and used exclusively for academic purposes in accordance with institutional ethical standards.

Figure 24. Positioning of the phone to record synchronized PCG and SCG data

SCG data contains many peaks associated with the aortic and mitral valve opening and closing. MC indicates mitral valve closing, AO means aortic valve opening, and RE is rapid ejection, the phase where blood is forcefully ejected. These three points are used in the model to separate individual heart beats to feed into the model and can be seen in Figure 25. A comparation between SCG and the corresponding ECG, characterized by points P-Q-R-S-T, is described in (Choudhary et al., 2019).

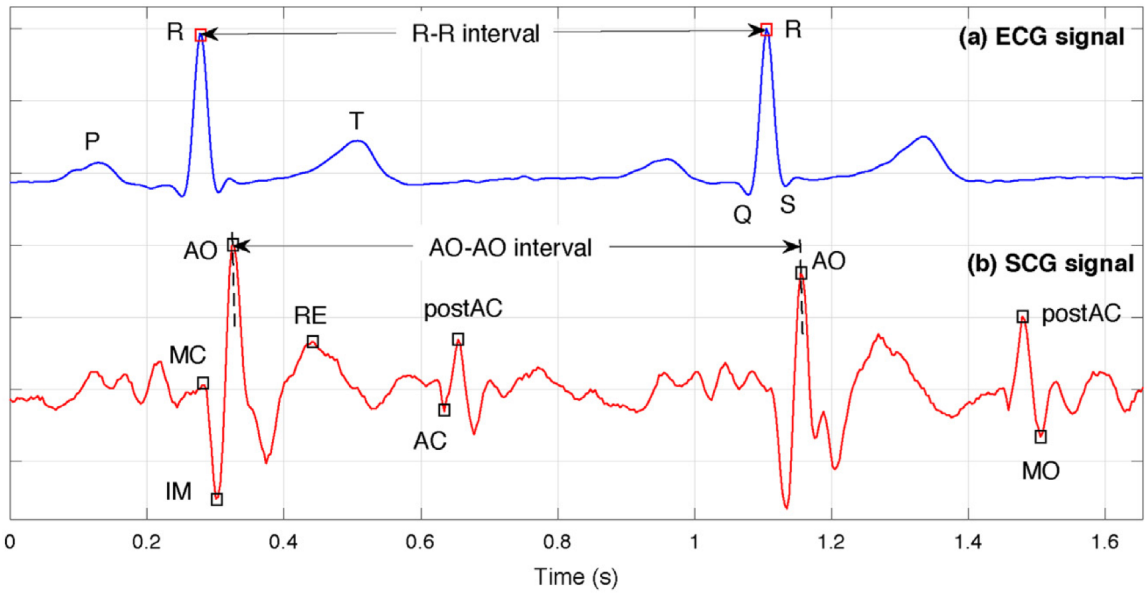

Figure 25. SCG signal in comparison to an ECG signal (Choudhary et al., 2019)

Another important heart signal is a phonocardiogram (PCG), which records acoustic signals from the heart (Figure 26). They can be recorded with a digital stethoscope in the hospital, or a microphone. Surprisingly, these sounds can be strong enough for a regular smartphone to capture and process, as well as a pair of in-ear headphones with microphone (such as Apple AirPods). Doctors listen to valve activity and blood flow turbulence which can provide information about heart murmurs for example. PCG data has 4 distinct heart sounds named S1-S4. All 4 are only seen in children and may suggest heart failure in adults if present. For the purpose of this project, and the subject the data was recorded on, S1 contains the most relevant sound information and it's well recorded.

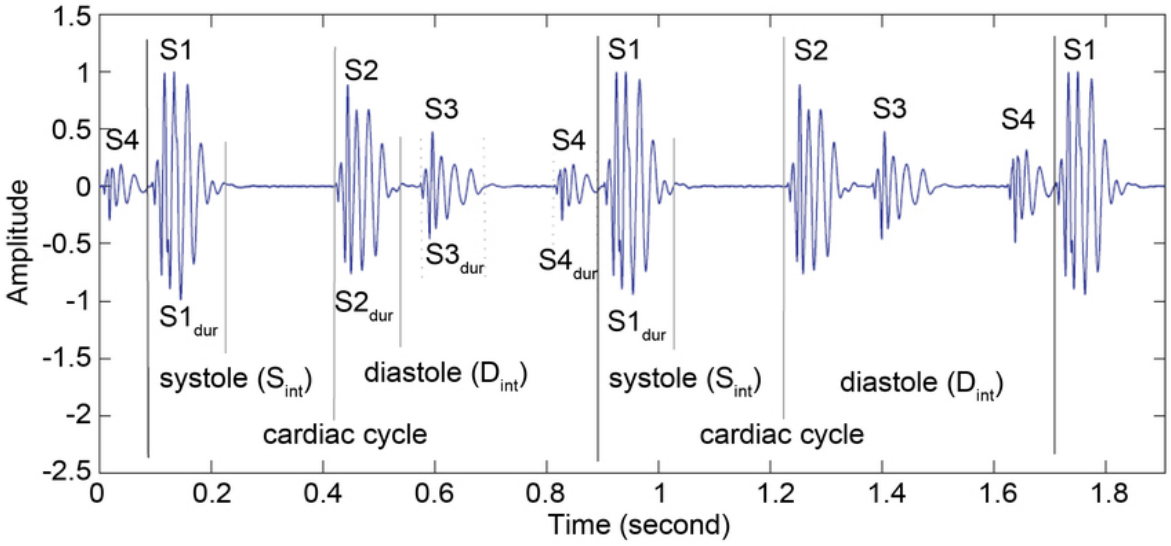

Figure 26. PCG signal (Nabih-Ali et al., 2017)

SCG data is more informative to doctors than PCG data, but it is harder to capture in the least invasive way possible. Because of this, a final model in this project was created to take PCG audio signal from in-ear headphones and process is to obtain the equivalent SCG reading. Given the complexity of the data, I chose to work with neural networks that could learn from these more intricate signals and map the recorded PCG data to SCG.

The first step is to obtain synchronized PCG and SCG signals. Using an android app developed in Justin Chan's lab at Carnegie Mellon University, both signals were captured on an Android phone placed on my chest right below my heart (as shown previously on Figure 24). Only one subject was used, and although over fifteen minutes of data were recorded, only about 1 minute was clean enough to accurately train and test the model.

To prepare the data for the model, several preprocessing steps were taken. The first step was to resample both signals to 1 kHz. This made it easier to process and work with. Next, the most noisy data points from the beginning and end of the recording were eliminated to keep the data as clean as possible. Afterwards, even though the signals were recorded simultaneously, there was a slight delay of 80 samples. To synchronize them, I tapped the phone during the recording creating peaks in both signals to align them properly.

After the signals were pre-processed, they were closely examined to verify they were properly recorded and resembled what real SCG and PCG signals taken in a hospital would look like, as seen in Figure 27, and Figure 28. The top signal shows what the real signal should look like, and the bottom one is the one recorded on the phone. Both the alienation of the peaks and the separation between signal markers were checked.

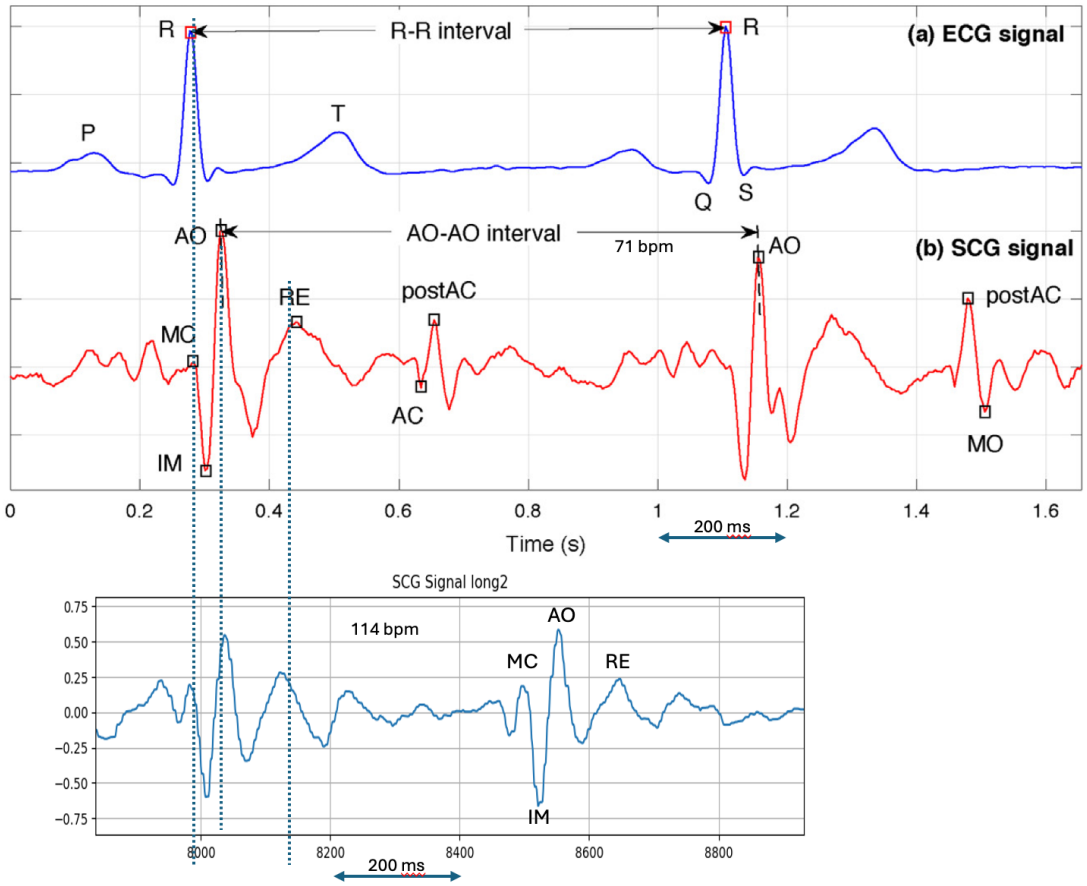

Figure 27. Verification of experimental SCG data (bottom) and theorical ECG/SCG signals (top). Top graph obtained from (Choudhary et al., 2019)

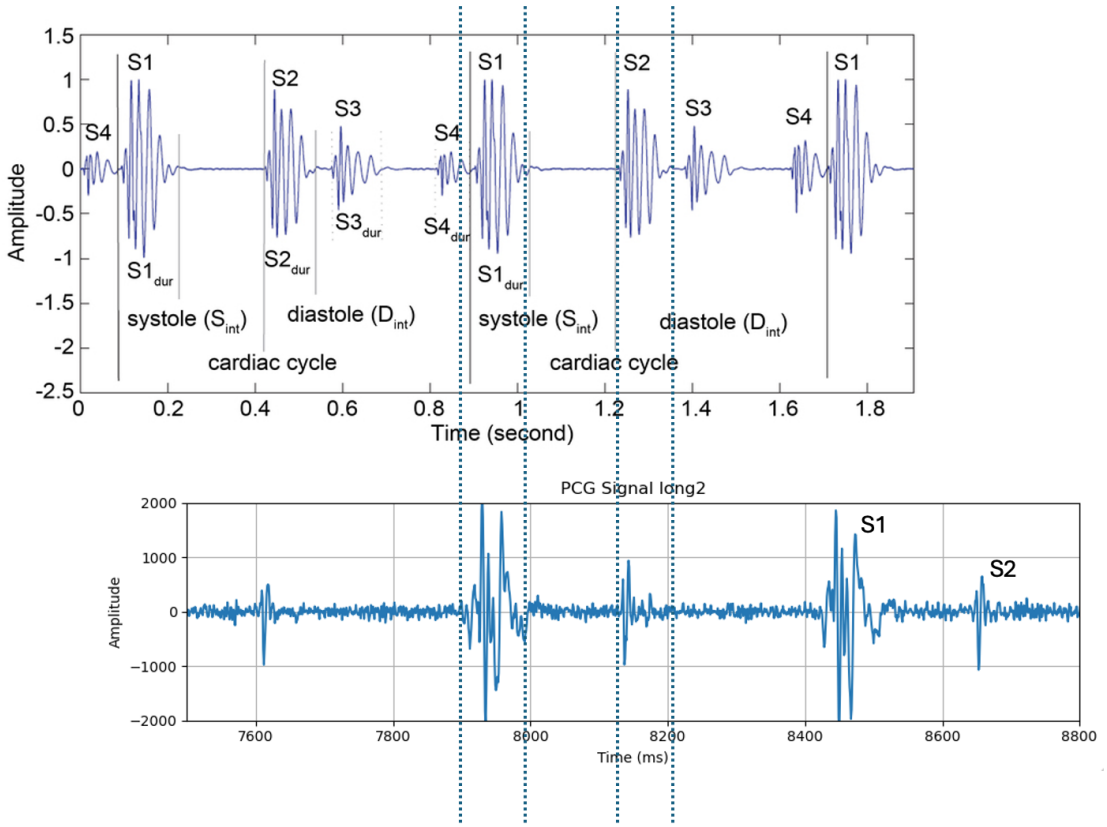

Figure 28. Verification of experimental PCG data(bottom) and theorical PCG signal (top). Top graph obtained from (Nabih-Ali et al., 2017)

Finally, the signals were cropped to only the sections containing heartbeats, as they are the most informative sections. To detect the heartbeats, the peaks of the SCG signal were used and then a margin on either side was set as the place to crop. This can also be done though the PCG signal. (Figure 29).

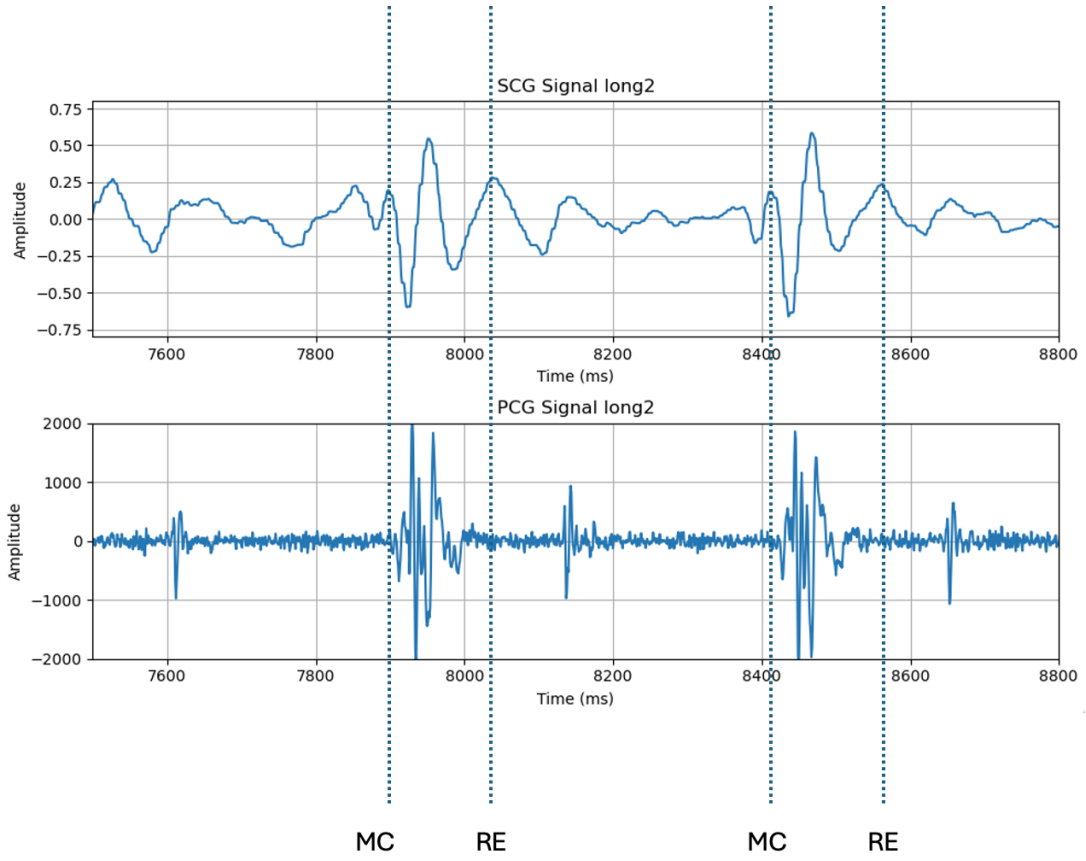

Figure 29. Plots of SCG and PCG signals, aligned and marked where they were cropped for the model

Now that the data is processed and cropped to the desired length, the problem being faced is a supervised learning problem where the input is a 1D signal (PCG, 135 samples), and the target/output is another 1D signal (SCG, also 135 samples). Therefore, this is a signal-to-signal regression problem. Several models have been developed which receive an input of PCG signal and predict a SCG signal as output. Three models taking different approaches were created and compared: CNN model, RNN model and transformer-based model.

The CNN model contains three convolutional layers which are fully connected, to transform the features from the input signal into the output signal. This model is ideal for capturing local patterns in the PCG data that correlate with SCG features.

The RNN model has a three-layer bidirectional Long Short-Term Memory (LSTM) model to process the signal sequentially. It aims to capture temporal dependencies across the signal and is good for tracking how patterns evolve over time in the cardiac cycle.

Finally, the transformer-based model was designed with a three-layer transformer encoder with multi-head attention. It uses a custom positional encoding for the time series, so it could capture both local and long-range dependencies in the signals. The result section will show that this transformer architecture is very accurate for this problem.

Chapter 6. RESULTS ANALYSIS

The following section presents the results obtained from the different models developed throughout this project. Each subsection focuses on a specific model, highlighting its performance and analyzing the results. The models are evaluated mainly using accuracy, balanced accuracy and loss. An effort was made to compare the performance of the proposed models with other alternative algorithms.

6.1 ACTIVE LEARNING MODEL ON WATER LOOP DATA

The active learning model was developed and tested on the two different types of data: slow pressure decreases and simulated kinking. Since the pressure sensors are positioned at the outlet of the circuit, they detect pressure drops when the flow is restricted, unlike locating the sensors between the pump and the obstruction, where the pressure increases while restricting the flow. The model was trained to compare the six different query selection methods to determine which one has the best balance in terms of accuracy and number of samples queried. Figure 30 shows the accuracy of each query method compared to the original data as a function of time. Throughout the training of the model, its predictions were accumulated and then a gaussian filter was applied to smooth the predictions and compare the accuracy of each method to what is happening in the data at that moment. In general, the metrics are very high, in part due to the imbalance of data. Figure 30 shows a zoom in of the slow pressure decline event. At the beginning of the graph (8800 s) there is a decrease in pressure which is quite significant and should be considered an anomaly. Later on, the pressure keeps dropping after 10200 s as a consequence of an important obstruction introduced in the circuit, and that is again a situation in which the model should keep detecting anomalies. Approximately below the dotted line shown in the graph as a reference, the data was marked as anomalous because it is around that area that the pressure drop becomes a problem. During in-vivo experiments, when the sheep changed position from

standing up to laying down or vice-versa, there were smaller changes in flow; as shown previously in Figure 22 the natural changes are smaller that 4%.

Out of all the query methods, the only one that barely drops in accuracy when the pressure declines: Stream probabilistic AL. This query method was able to select the most informative data points as inputs and correctly detect the anomalies. The rest of the methods were not as quick to detect the anomaly. Stream random sampling does the worst job as its accuracy keeps decreasing at random points where all the other models are able to predict accurately.

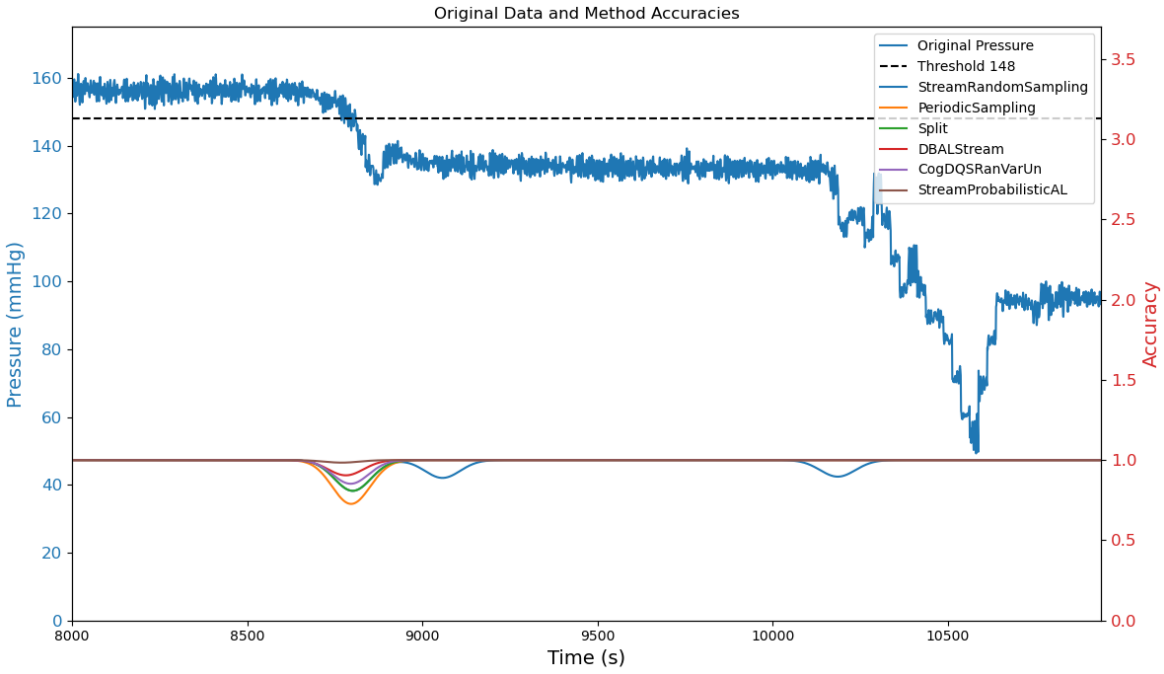

Figure 30. Accuracies of each query method during the recording

Table 1 shows the different metrics calculated for each query method. Unsurprisingly all are above 95% as seen in the graph above. The data was simple and easy to predict in this laboratory experiment. An interesting aspect to investigate is the acquisition count. Previously, it was noted that Stream probabilistic AL was the best query strategy as it didn't decrease in accuracy when the pressure dropped because and induced anomaly. However, it has the highest sampling rate which means it would not be optimal for this task as it can't be querying all the data points. This would defeat the purpose of the model. Based on Figure

30, the second-best model in accuracy is DB AL Stream. It has good accuracy, and based on Table 1, also a low acquisition count, which together mean that this would be a good candite for a final model. Furthermore, the third best model based on accuracy is CogDQRanVarUn, and it has the lowest sampling rate. This would be another optimal candidate to consider as it has the most reasonable acquisition count and still maintains a high accuracy.

Table 1. Metrics calculated for each query method (the closer to 1, the better) and the acquisition count for the slow descent in pressure

Query strategy	Average accuracy	Balanced accuracy	Sensitivity	Specificity	Acquisition Count
Stram Random Sampling	0.9952	0.9879	0.9757	1	539
Periodic Sampling	0.9966	0.9959	0.9949	0.997	544
Split	0.9977	0.9942	0.9883	1	590
DB AL Stream	0.9989	0.999	0.9991	0.9989	545
CogDQRanVarUn	0.9983	0.9963	0.993	0.9995	308
Stream Probabilistic AL	0.9998	0.9997	0.9995	0.9999	10888

While the tables in this Results section widely show excellent results, the graphs really help isolate the class imbalance issues that can also be seen in the balanced accuracy measurement.

The same analytical procedure was carried out for the data obtained during a simulated kinking event, which produced an abrupt drop in pressure. Figure 31 shows the accuracies of each query method throughout the complete data recording. At first glance, stream random AL is again the worst method as its accuracy drops to around 50% during the kinking, while all other methods are able to detect the anomaly during the time of pipe obstruction.

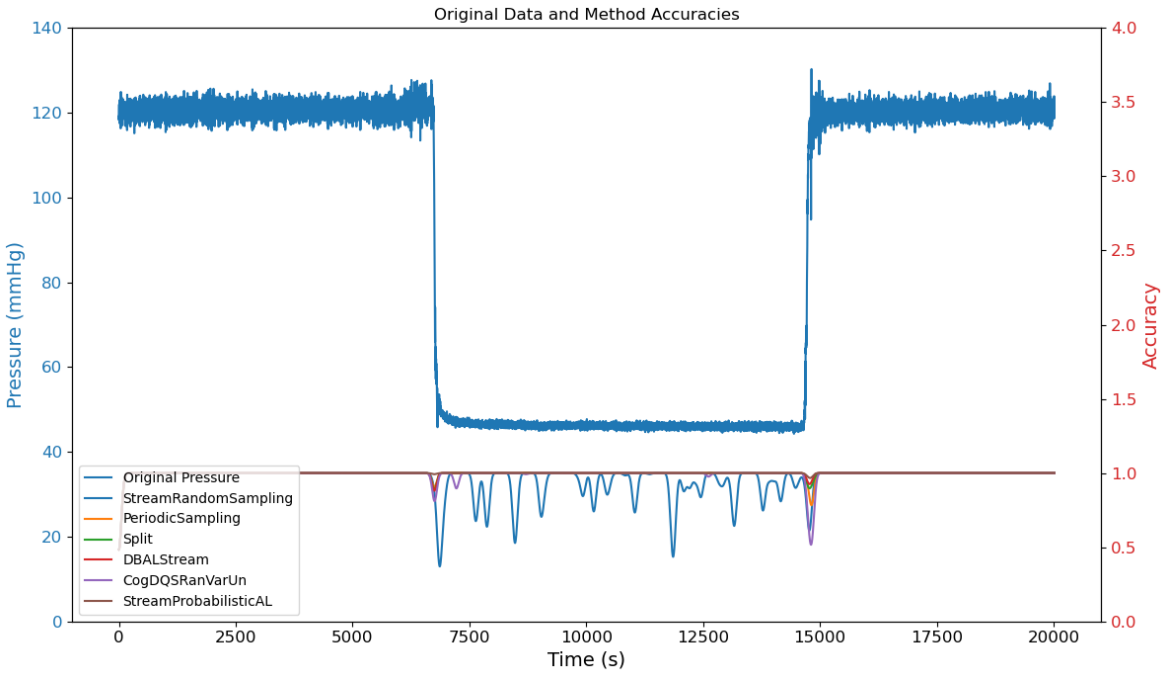

Figure 31. Accuracies of each query method during the recording on kinking data

To have a better understanding of how a sudden pressure change affects the accuracy of each method, Figure 32 and Figure 33 show a detailed view of those areas. In both cases, stream probabilistic active learning has the best accuracy, with values very close to perfect 1 during all the experiment, meaning that the model can correctly predict normal condition or anomaly during the whole experiment. However, in this case, the accuracy of CogDQSRanVarUn method drops to the second lowest value during pressure decrease and it is the worst during pressure recovery, while in the previous experiment this method was the third best in accuracy. This method did worse in the kinking experiment because the pressure change was very sudden and just by sampling data points at random, it is not capable of learning at the necessary rate. In the other experiment, even with random sampling, due to the slow pressure decrease, the model is able to detect the change.

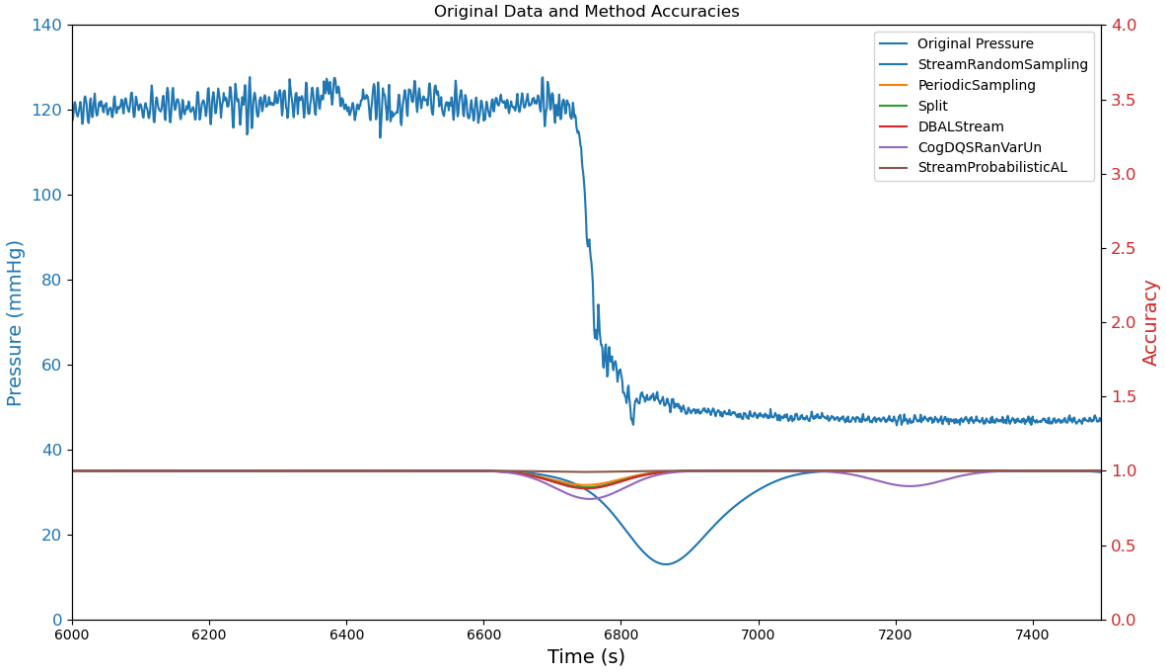

Figure 32. Zoom in to the sudden pressure drop comparing the real data with the accuracies of the different query methods

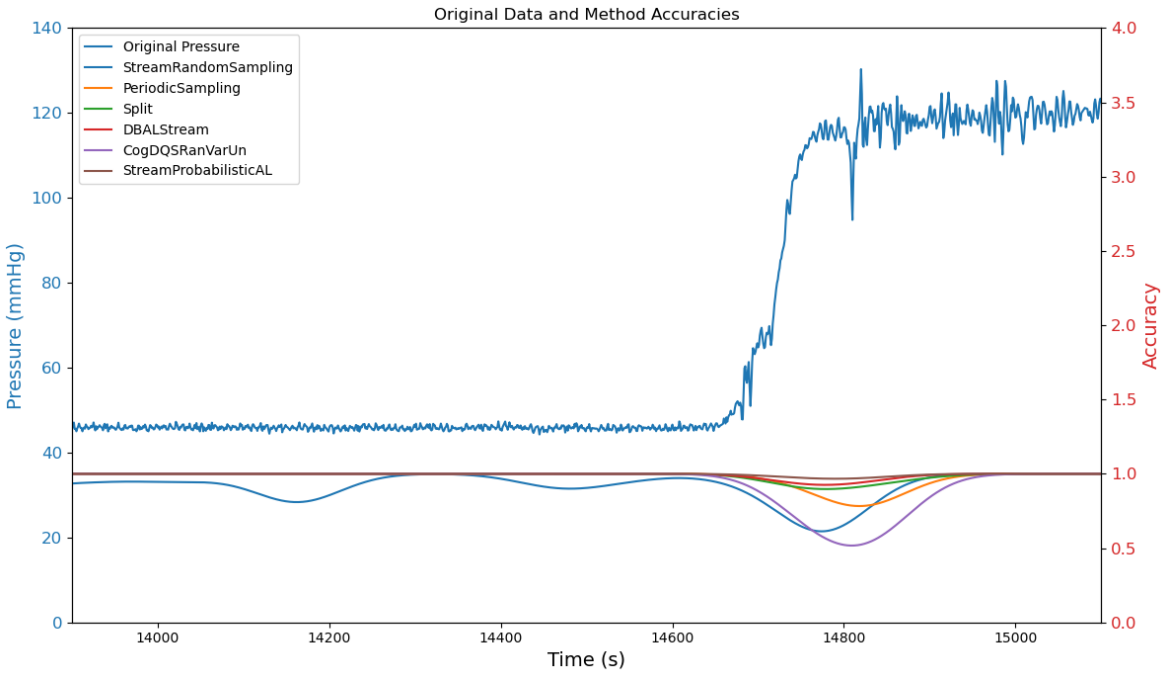

Figure 33. Zoom in to the sudden increase comparing the real data with the accuracies of the different query methods

Finally, the metrics are shown in Table 2 and as it was the case before, stream probabilistic AL has the best accuracy but also the highest acquisition rate, at double or 4 times more samples than the other methods. CogDQSRanVarUn has the lowest acquisition count again, but this time it came with the worst predictions in the key moments of pressure change. Overall, the worst method is Stream random sampling.

Table 2. Metrics (the closer to 1, the better) and the acquisition count calculated for each query method for the kinking data

Query strategy	Average accuracy	Balanced accuracy	Sensitivity	Specificity	Acquisition Count
Stram Random Sampling	0.9675	0.9597	0.9205	0.9989	1055
Periodic Sampling	0.9979	0.9979	0.9979	0.9979	999
Split	0.9988	0.9985	0.9971	0.9999	1080
DB AL Stream	0.9987	0.9985	0.9975	0.9995	1000
CogDQRanVarUn	0.9963	0.9962	0.9959	0.9966	557
Stream Probabilistic AL	0.9998	0.9998	0.9996	0.9999	19983

Split and **DB AL Stream** are selected as the best methods as they show a good balance in high accuracy and average Acquisition Count. For both methods, sensitivity and specificity

show high values which indicates that the model is able to accurately predict for both classes. Balanced accuracy and accuracy are also both relatively high (only surpassed by Stream Probabilistic AL) which also supports that the imbalance of data does not significantly affect the model.

In active learning, it is important to not only look at accuracy but also acquisition rate as it shows how many resources it's using and if it is a better option that traditional machine learning.

6.2 ACTIVE LEARNING MODEL ON SHEEP DATA

After checking the functionality of the model on the in-vitro data, the next step was to validate active learning on in-vivo data. A portable ECMO device along with several sensor was installed on a sheep to run experiments while collecting in-vivo data. The data from the first experiment done on the sheep contained several suction events where the flow suddenly dropped and consequently the pressure suddenly increased. Because they were suction events that primarily affected the flow rate, a model was trained first on just flow data and then modified to take both flow and pressure as inputs. Using several input signals allows the model to make more advanced decisions, such as if the flow drops below a give dynamic threshold and at the same time the pressure rises above another dynamic value, then we have an anomaly.

6.2.1 FLOW MODEL

The first step taken in this model was to check and compare different numbers of training points using pool-based active learning and additionally comparing different traditional machine learning models. The query method being used was random sampling, which means that there is no specific algorithm behind deciding which samples to query. It just selects some of the input samples at random.

The different number of training points indicate how many instances the model can randomly select to train. Clearly, with more training samples, the models performed better while typically become slower detecting sudden anomalies. The accuracies of each machine learning method and for different number of training points (2, 10, and 20) are shown in Table 3. Out of these results, decision tree was chosen as the desired algorithm as it is simple and easy to explain yet obtained high accuracy using only 20 samples, while outperforming other machine learning algorithms.

Table 3. Accuracy of several algorithms for different number or input points: 2, 10, 20 (the closer to 1, the better)

Method	2 TP	10 TP	20 TP
SVM	0. 6008	0. 6008	0. 6008
Gaussian Process 1.0 * RBF(1.0)	0. 6008	0. 6182	0. 5905
Decision Tree	0. 6008	0. 9014	0. 9014
Decision Tree, max_depth=5	0. 6008	0. 9014	0. 9014
Random Forest	0. 6008	1	1
Random Forest, max_depth=5, n_estimators=10	0. 6008	1	1
Neural Net	0. 5839	0. 6006	0. 5999
Neural Net, alpha=1, max_iter=1000	0. 6008	0. 6008	0. 6008
Naive Bayes	0. 6008	0. 6695	0. 7210
QDA	0. 6008	0. 6610	0. 9295
Logistic Regression	0. 6008	0. 6281	0. 6304

Once the model was chosen, the same process as with in-vitro data was followed to compare different query methods in the active learning approach. All six methods were trained and the results, in a similar format as before are shown in Figure 34. The figure shows a close up of a suction event where, as in the case of in-vitro testing, the best query method is Stream Probabilistic.

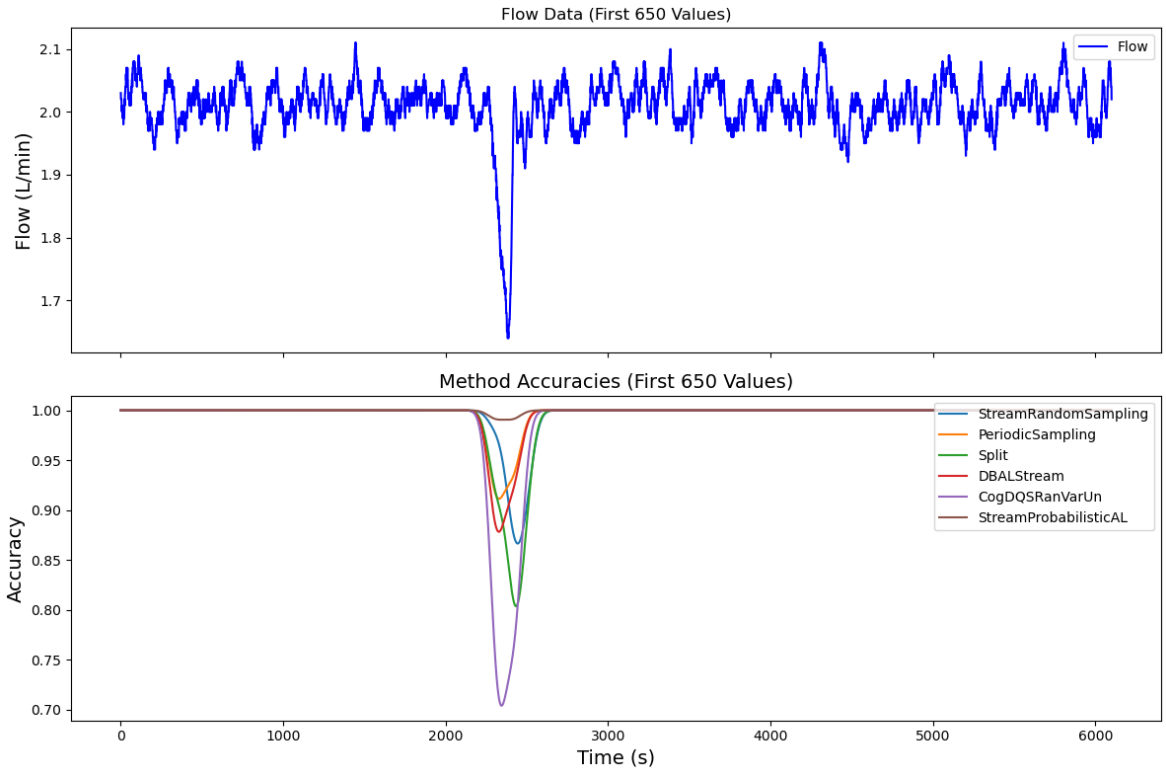

Figure 34. Comparison between flow data on a suction event with the accuracies of different query methods

Table 4 shows a comparison of the different method's metrics. stream probabilistic active learning has the highest acquisition rate, over 10 times as much as other methods. It queries almost all the data samples. This means that although the model may barely predict incorrectly, its high acquisition count shows that it is not a viable solution for the current problem. Other methods achieve high accuracies with much lower acquisition counts. In this situation, **Periodic Sampling** and **DB AL Stream** have the lowest sampling rate while still maintaining a high accuracy.

Table 4. Compilation of the metrics for each query method (the closer to 1, the better) as well as the number of points queried for the model analyzing flow

Query strategy	Average accuracy	Balanced accuracy	Sensitivity	Specificity	Acquisition Count
Stram Random Sampling	0.9957	0.9008	0.8034	0.9983	9554
Periodic Sampling	0.9969	0.9319	0.8652	0.9986	9531
Split	0.9966	0.9439	0.8898	0.9981	10252
DB AL Stream	0.9969	0.9265	0.8541	0.9989	9531
CogDQRanVarUn	0.9947	0.906	0.8149	0.9971	4640
Stream Probabilistic AL	0.9998	0.9946	0.9893	0.9999	190629

Contrary to in-vitro data, in this case, sensitivity across the board tends to be lower or significantly lower that specificity. This means that there are few false positives, but many false negatives. As seen in Figure 34, the accuracy notably decreases during suction events which means that there are many false negatives (the model is predicting normal instead of anomaly). Finally, balanced accuracy also shows a drop compared to normal accuracy, which means that the model is better at detecting normal data than anomalies. However, since the gap is not always significant, this imbalance does not affect the model.

6.2.2 FLOW AND PRESSURE MODEL

In an effort to improve the previous model, a new variable was collected and added as an additional input to the model. This would increase the input data and give the model more information to learn from. Figure 35 shows both flow rate in blue and pressure in red, in addition to the accuracies for each query method. As was the case in the previous models, the accuracy decreases on the suction events. In the zoom-in graph shown in Figure 36, it's very clear how much each method decreases in accuracy near the event. Stream Probabilistic AL is again the best at predicting and CogDQSRanVarUn's accuracy reduces significantly. However, Stream Probabilistic sampling has the best maintained accuracy throughout the suction event.

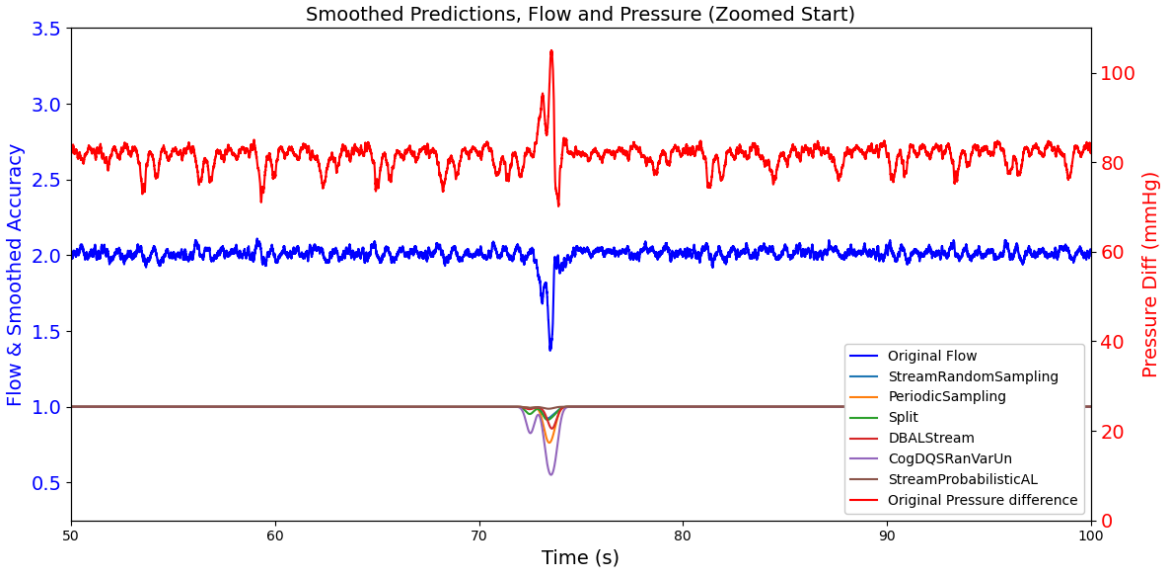

Figure 35. Predictions of all query methods on a suction event (bottom plot) and showing the flow rate (L/min) in blue and the pressure in red

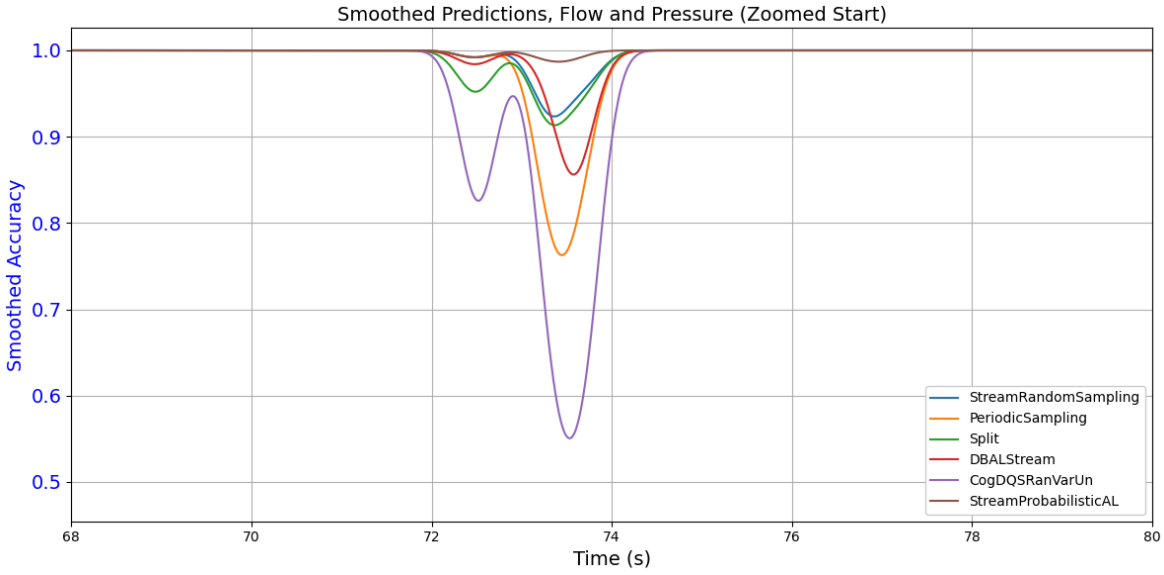

Figure 36. Zoom in of accuracy during suction event

Table 5 shows the comparison of the metrics for each query method. As was true in the previous cases, Stream Probabilistic AL has the highest acquisition rate, making it a less likely candidate compared to other methods. After discarding Stream Probabilistic AL, the

next best accuracy is from **Split**, which additionally has a more reasonable acquisition count and its accuracy doesn't significantly drop in suction events.

Table 5. Compilation of the metrics for each query method (the closer to 1, the better) as well as the number of points queried for the model analyzing both flow and pressure

Query strategy	Average accuracy	Balanced accuracy	Sensitivity	Specificity	Acquisition Count
Stram Random Sampling	0.9963	0.9187	0.8391	0.9984	9518
Periodic Sampling	0.9968	0.9397	0.8811	0.9983	9551
Split	0.9974	0.9474	0.8962	0.9987	10277
DB AL Stream	0.997	0.938	0.8775	0.9986	9552
CogDQRanVarUn	0.9946	0.9089	0.8208	0.9969	4654
Stream Probabilistic AL	0.9997	0.9928	0.9857	0.9998	191029

In this case, sensitivity and specificity are sometimes significantly different, with a decrease of 10%. This indicates that the model's accuracy decreases during suction events. However, comparing these values with the model with only one variable (Table 4), this performs 2-3% better across all the metrics in the board, showing that the model makes better decisions when it has access to more information.

6.3 BREATHING DETECTION MODEL

From the last sheep experiment, the first two days of clean data exhibited a distinct pattern that turned out to be the animal's breathing rate. To determine the breathing rate over time, a sliding window of two minutes was used to capture an accurate value of the breathing rate. Two different methods were employed to compare: peak detection in the time domain and FFT in the frequency domain.

Starting simple, the first model implements a peak detection algorithm to count the number of breath events in a given time windows. For each time window, the model detects the peaks (Figure 37), counts the number of peaks and divides by the time difference (seconds) from the first peak to the last. After computing the breathing rate in this way it is multiplied by 60, to obtain the breathing rate of the person, or sheep in this case, in breaths per minute.

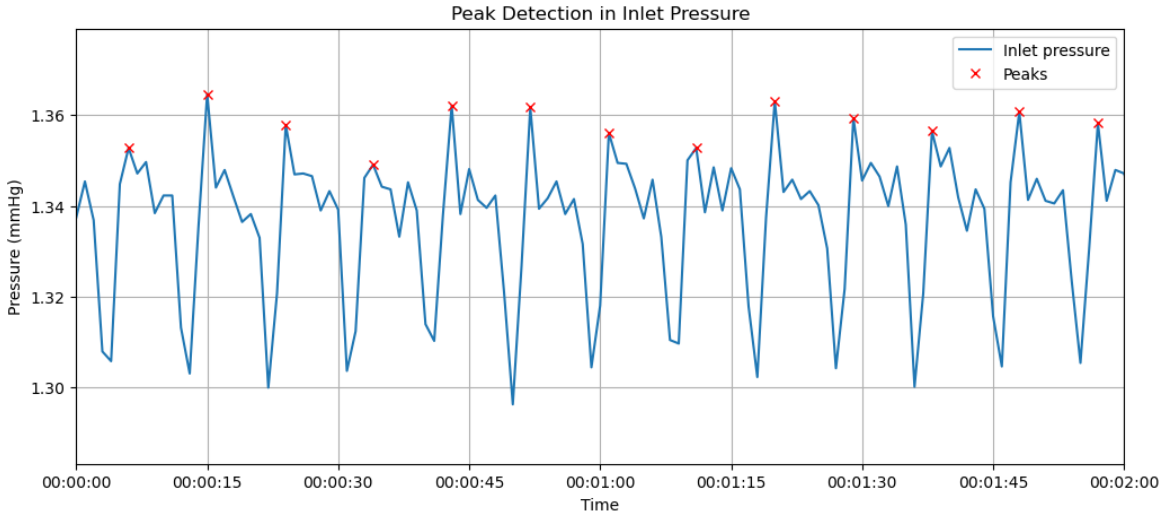

Figure 37. Detected peaks on inlet pressure sheep data

To visualize the results, a moving average was applied to smooth the results and an average of 6.5 breaths per minute can be seen throughout the recording. (Figure 38).

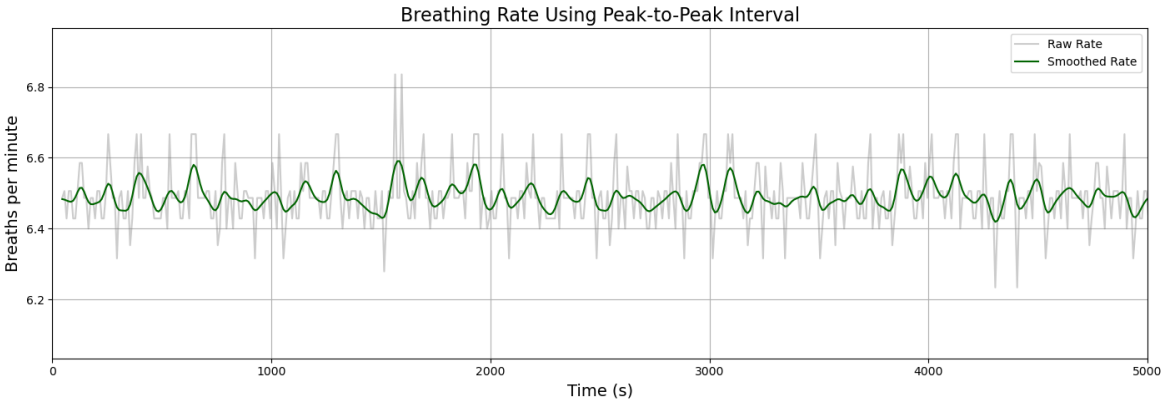

Figure 38. Smoothed breathing rate over the experiment

To verify the accuracy of the method without any ground truth, the breathing rate was calculated by hand on several segments of the sheep data by simply counting the number of peaks in a one-minute window. This can be seen in Figure 37, where from 19:27:30 to 19:28:30, there are 6 peaks, almost 7, which supports the results of 6.5 breaths per minute. The average respiratory rate of a sheep is between 12 and 20 breaths per minute which at first glance make these results seem incorrect. However, when patients use ECMO machines

their breathing rate is greatly reduced because the blood going into the lungs is already oxygenated in the machine (Abrams et al., 2020).

As an alternative method, the breathing rate was calculated using a Fourier transform to detect the main frequency of the signal. Figure 39 shows an FFT of a 2 minute window, where the first peak is clearly the breath rate at just over 6 breaths per minute (0.1 Hz = 6 breaths/min).

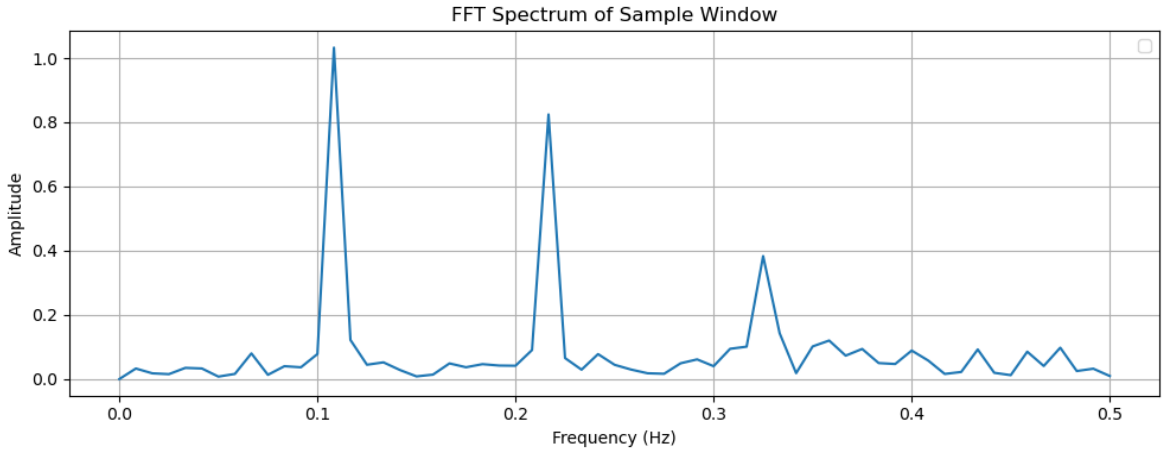

Figure 39. FFT of a segment of the inlet pressure data with the breathing pattern

Therefore, to make the FFT model work in real time, it was set up with a sliding window of size 120 samples, the same conditions as the peak detection algorithm. It produced similar results of a stable rate as seen in Figure 40 . Note that in the case of using FFT with a windows size of 120 s, the resolution in the frequency domain is 1/120 which corresponds to 0.008333 Hz or 0.5 breaths per minute. So, the value obtained for the sheep is 6.5 b/min and may only jump to 6.0 or 7.0 since no intermediate values are not possible.

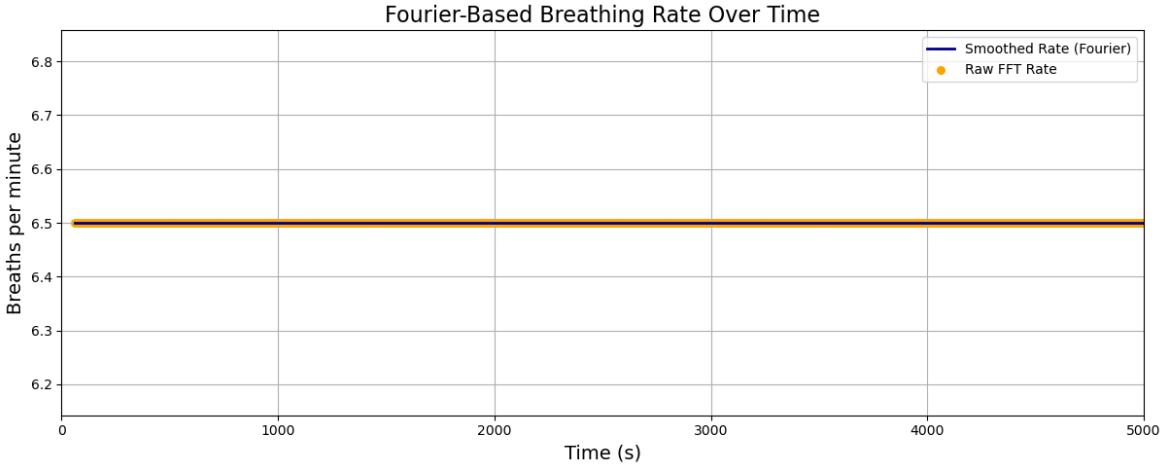

Figure 40. FFT breathing rate detected throughout the recording, zoomed in to the beginning

While frequency-domain methods like FFT can provide a general estimate of breath rate (or other biomedical signals), they are often less accurate than time-domain approaches such as peak detection, particularly in short or noisy signals. FFT offers a smoothed, averaged view of periodic components and lacks beat-level resolution. In contrast, peak detection directly measures inter-breath intervals, enabling precise breath rate calculations and capturing variability that FFT cannot reflect. As supported by (Skotte & Kristiansen, 2014), time-domain peak detection remains the preferred method for accurate breath rate estimation in many biomedical applications.

6.4 HEART SIGNAL DETECTION MODEL

The objective of this last model is to monitor the patient's heart rate by generating their seismocardiogram (SCG) signal from a phonocardiogram (PCG) recording. As described in detail in Chapter 5, the SCG signal has a direct correspondence to the ECG, and the doctors are trained to analyze such SCG signal. SCG signal is very difficult to measure with a smartphone without noise, however, the PCG signal can be easily recorder with headphones equipped with microphones such as AirPods. Therefore, the interest to be able to obtain an SCG signal out of a PCG signal.

The data used to train the model were collected from a single subject—myself—gathering synchronized phonocardiogram (PCG) and seismocardiogram (SCG) signals. These signals were recorded simultaneously using a smartphone placed on my chest, near the heart, under voluntary and informed consent. All data collection was conducted with full awareness of the research purpose, and no personally identifiable information was stored or shared. All the data was managed securely in accordance with ethical guidelines for personal data handling. These resulting signals were then aligned and correspond beat by beat as seen in Figure 41.

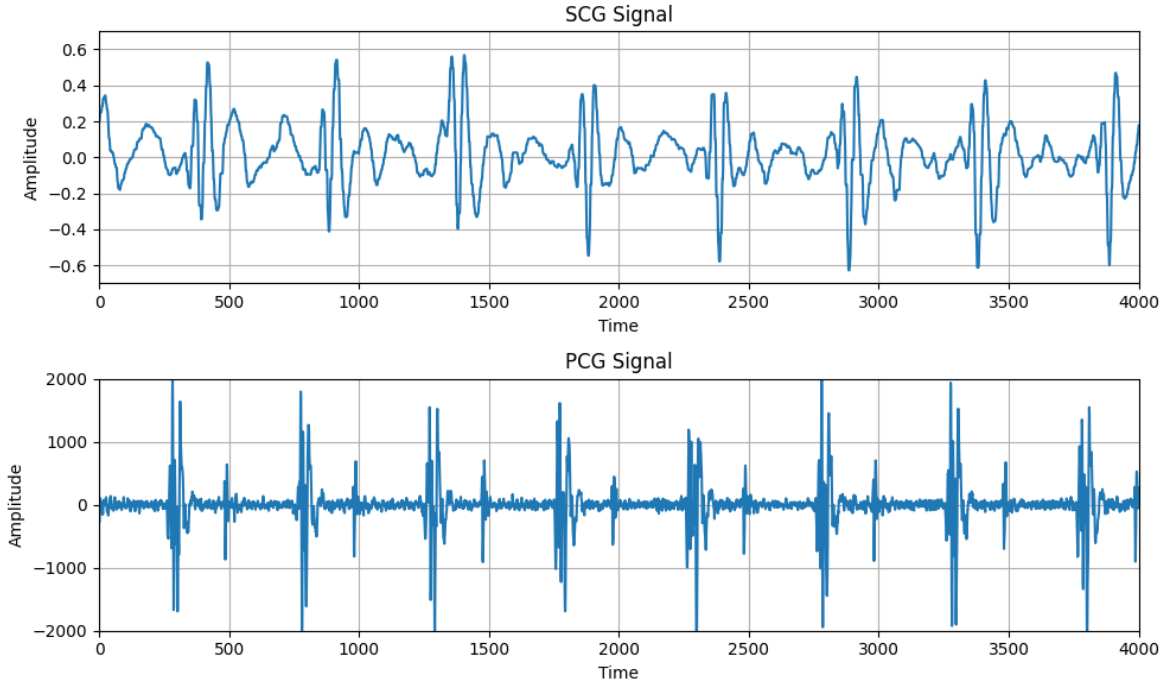

Figure 41. Comparison between SCG data and PCG data recorded at the same time

In order to train the model, I used only the sections of the signal corresponding to heart beats, which contain the majority of the information. To automatically crop the signals, the center peaks of the SCG heart beats were detected using a simple peak algorithm (Figure 42). Then, the two minor peaks were calculated from the central one to delimit the beginning and end of the signal to be used, as shown in Figure 43. The same crop was applied to both signals

and used as input for the model. Each input pair contains a cropped heartbeat from both PCG and SCG.

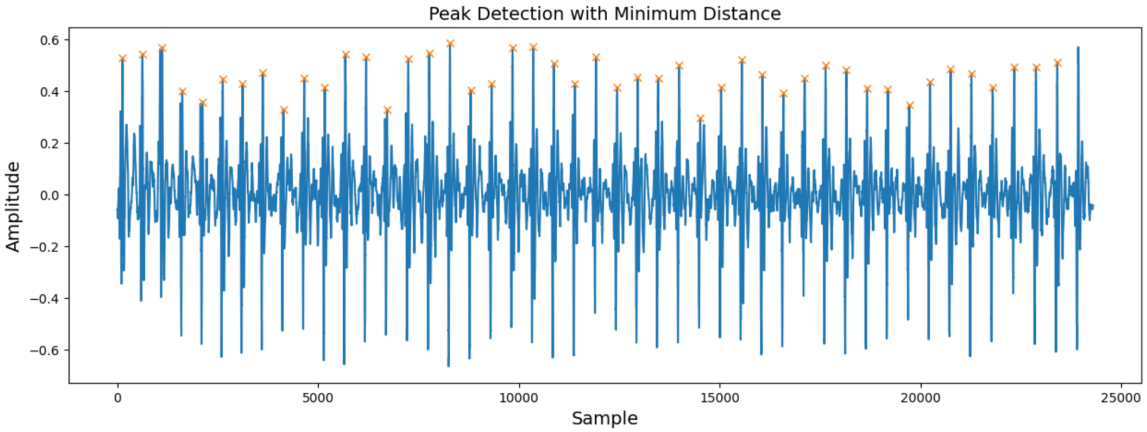

Figure 42. Peak detection done on SCG data to crop the informative sections

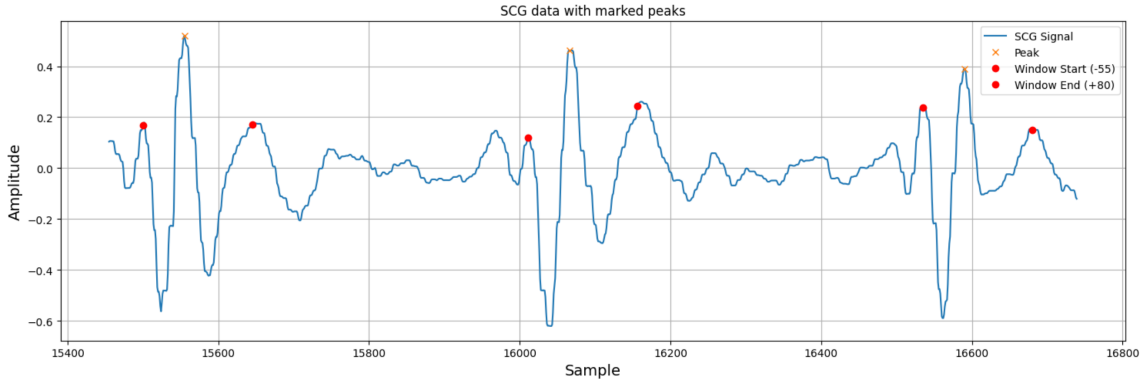

Figure 43. Close up of peak detection, with three main peaks being identified

Once the data is preprocessed and cropped, three different models were trained and compared. Any given PCG heartbeat, the output of the model is only the corresponding SCG heartbeat, not the signal in between heartbeats. To reconstruct the output, the distance between heartbeats is determined from the PCG signal and in between it is filled with zeroes. The signal in that section is not valuable. The key parts are the heartbeat and the distance between them. This will be shown below in Figure 44, Figure 47, and Figure 50.

The performance of each model was judged but simply comparing the predicted vs the original data, and with R^2 score. This coefficient of determination measures how well your model explains the variability of the target data. It is defined as:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

Where SS_{tot} is the total sum of squares (variance of actual data from the mean) and SS_{res} is the residual sum of squares from the variance of actual values and predictions. The closer this value is to 1, the closer the model is to perfect predictions. An R^2 score of 0 means that the model might as well predict the mean value of the dataset. Finaly, a negative value indicates that the model is predicting worse than the average.

As described in Chapter 5, 3 different models were trained to compare results and select the best approach for SCG signal estimation. The following subsections show the results of these models.

6.4.1 CNN MODEL

The first model, was a convolutional neural network containing three convolutional layers and fully connected layers. The idea was that this structure would be able to capture local patterns in the PCG data and correlate them with SCG features. However, this model was unable to learn well from the data and its predictions don't match the expected SCG values. Figure 44 shows on the top graph the original SCG data in blue (the target signal) and the reconstructed CNN's prediction (model output) on top in orange. Below is the corresponding PCG signal (model input), it is very clear that the predicted signal is not correct.

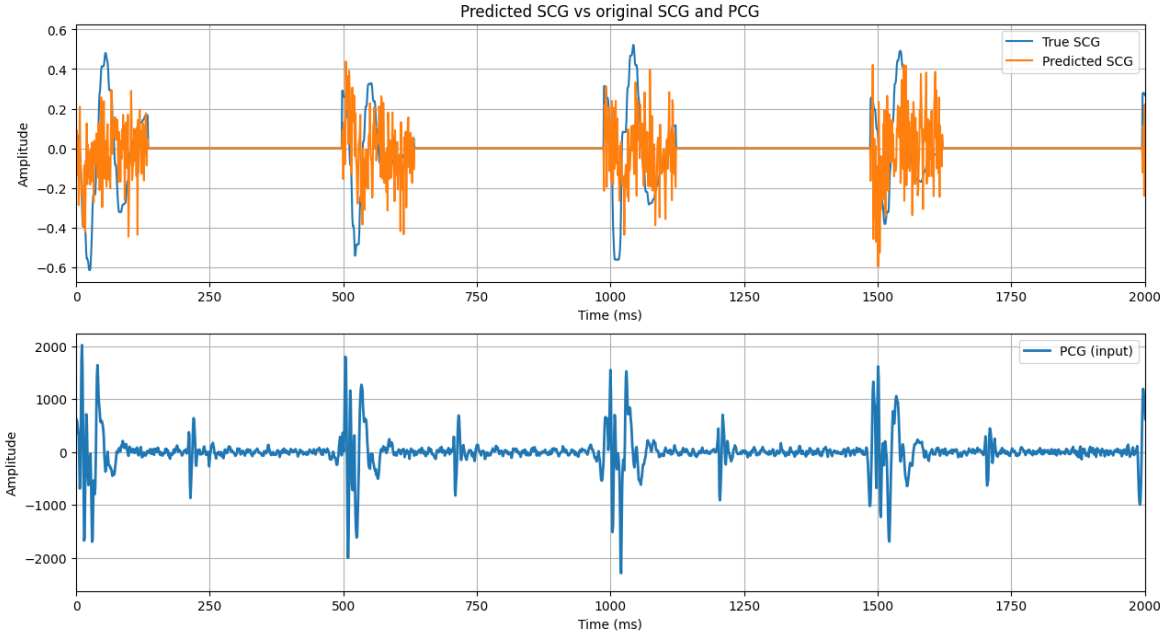

Figure 44. Results from CNN, comparison between original and predicted SCG (top) with corresponding PCG (bottom)

As seen in Figure 45, the loss drops quickly in the first 10 epochs but afterwards the model shows no improvement. Throughout training, validation loss remained consistently lower than training loss. This behavior is expected due to the presence of regularization (e.g., dropout) applied during training but not during evaluation. As a result, training loss includes additional penalty terms or noise, while validation loss reflects the model's raw performance on unseen data.

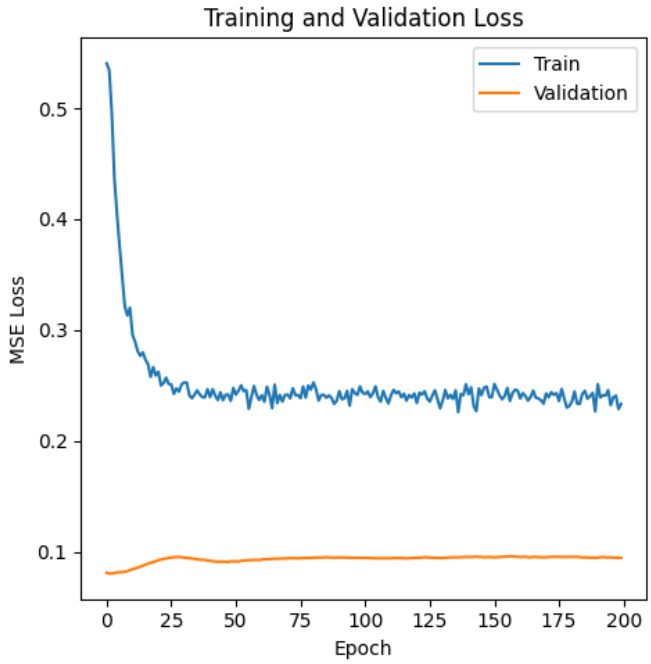

Figure 45. Results from CNN, train and validation loss throughout the epochs

The validation R^2 thought the epochs shows terrible results (Figure 46). It is consistently negative which means that the model is constantly predicting worse than the average signal value.

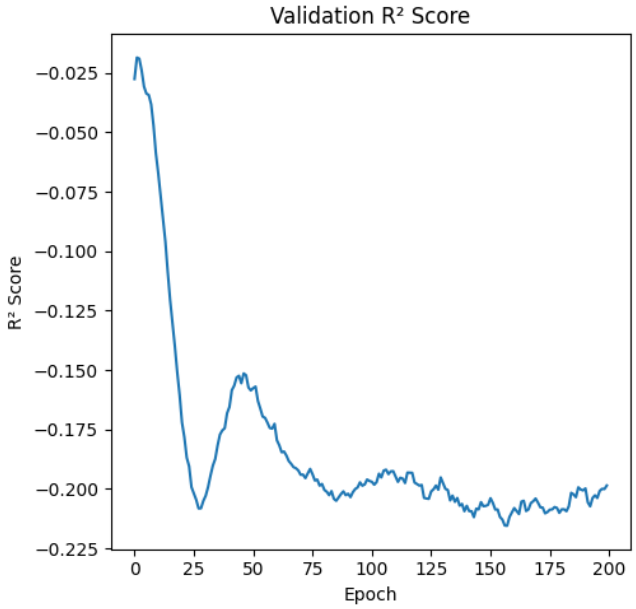

Figure 46. Results from CNN, validation R2 score throughout the epochs

6.4.2 RNN MODEL

The second model was a recurrent neural network (RNN) based on gated recurrent units (GRUs). Its objective was to capture temporal dependencies in the PCG signal and use them to reconstruct the corresponding SCG waveform. In contrast to CNNs, which process input data in a spatial manner, RNNs process signals sequentially, a property that renders them particularly well-suited for time-series prediction tasks. However, despite its theoretical advantages for temporal modeling, this RNN also failed to produce accurate SCG reconstructions, although it does significantly better than the CNN. As illustrated in Figure 47, the top plot presents the ground truth SCG signal in blue, with the RNN's prediction superimposed in orange. The lower plot displays the input PCG signal. The predicted signal does not fully align with the expected shape or timing of the SCG waveform, indicating that the model did not effectively learn the underlying mapping.

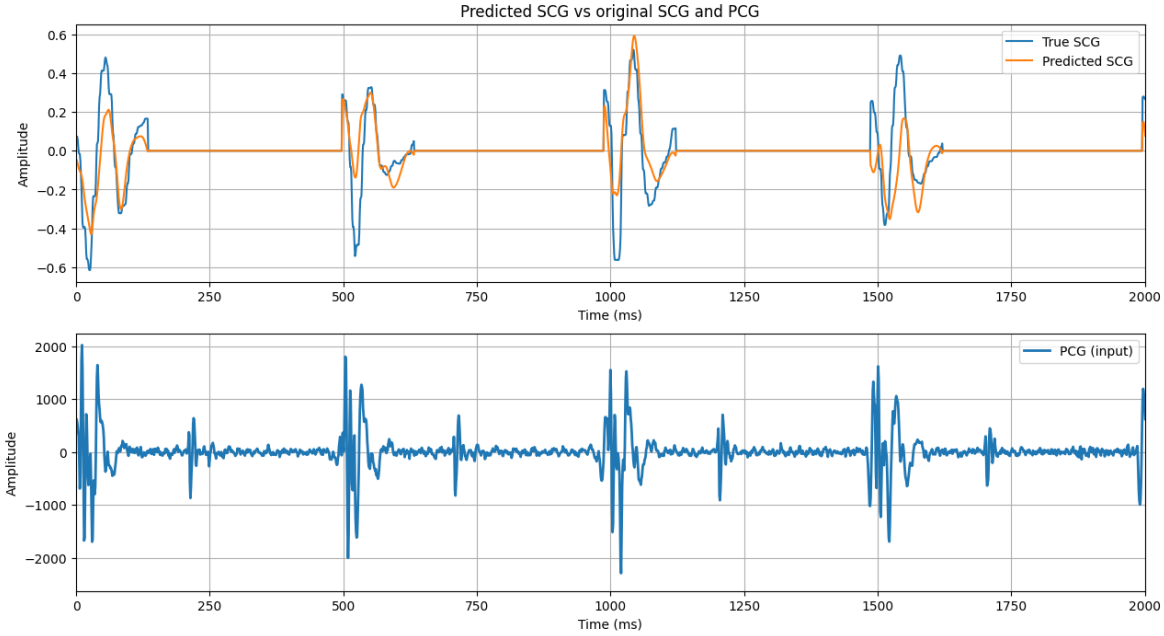

Figure 47. Results from RNN, comparison between original and predicted SCG (top) with corresponding PCG (bottom)

In this case, both validation and train loss decrease significantly throughout the first 100 epochs but then stabilizes and doesn't learn any new information (Figure 48). Compared to the CNN, the loss in this model is significantly less that before. In the previous model it began at 0.5 and ended in 0.25. Here, the loss begins lower than the previous best training loss.

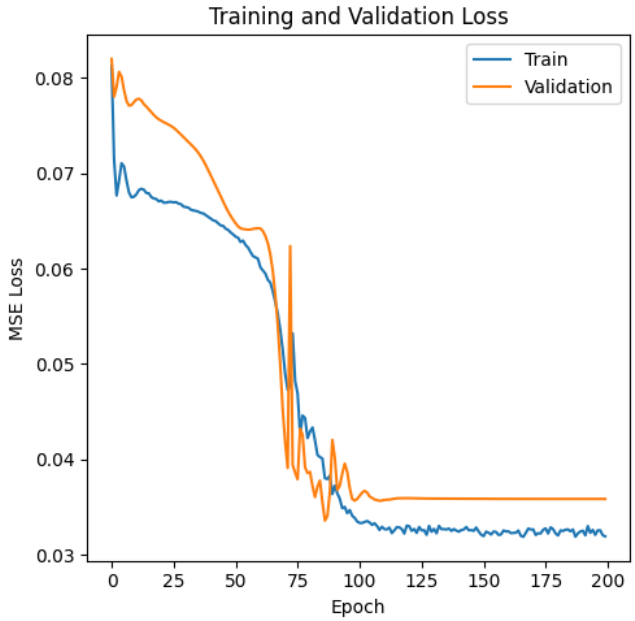

Figure 48. Results from RNN, train and validation loss throughout the epochs

Another improvement in this model is the R^2. Although it starts negative, it progressively increases during the first 100 epochs, until the model stops learning. However, the R^2 is at around 0.5 which is not an optimal value. There is still room for improvement.

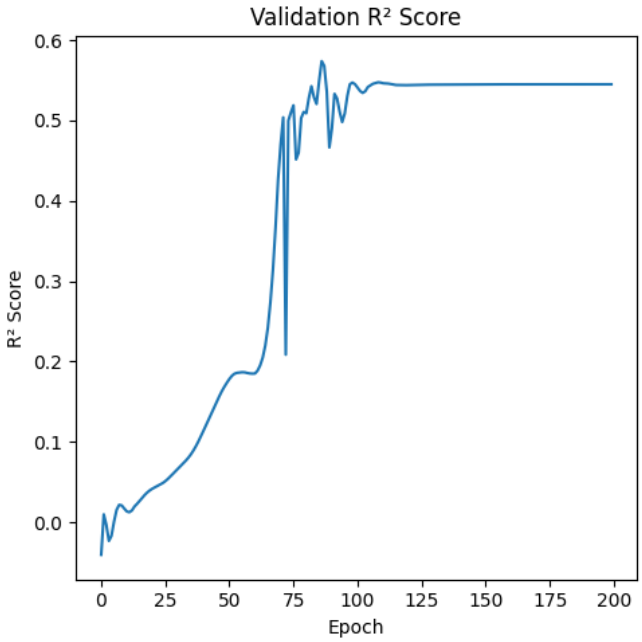

Figure 49. Results from RNN, validation R2 score throughout the epochs

In conclusion, this model's predictions are significantly better than the CNN but there is still room for improvement. For this reason, a last model was trained using transformers.

6.4.3 TRANSFORMERS MODEL

The third and last model implemented was a transformer-based architecture, leveraging self-attention mechanisms to model long-range dependencies within the PCG signal. This architecture is particularly powerful for capturing global patterns and aligning relevant features across time steps, which made it a promising candidate for SCG reconstruction. After training, the Transformer model demonstrated superior performance in capturing the general structure of the SCG waveform when compared to both the CNN and RNN. It not only captures the shape of the peaks and the proportional amplitudes between them, but it also matches the expected amplitudes quite closely (Figure 50).

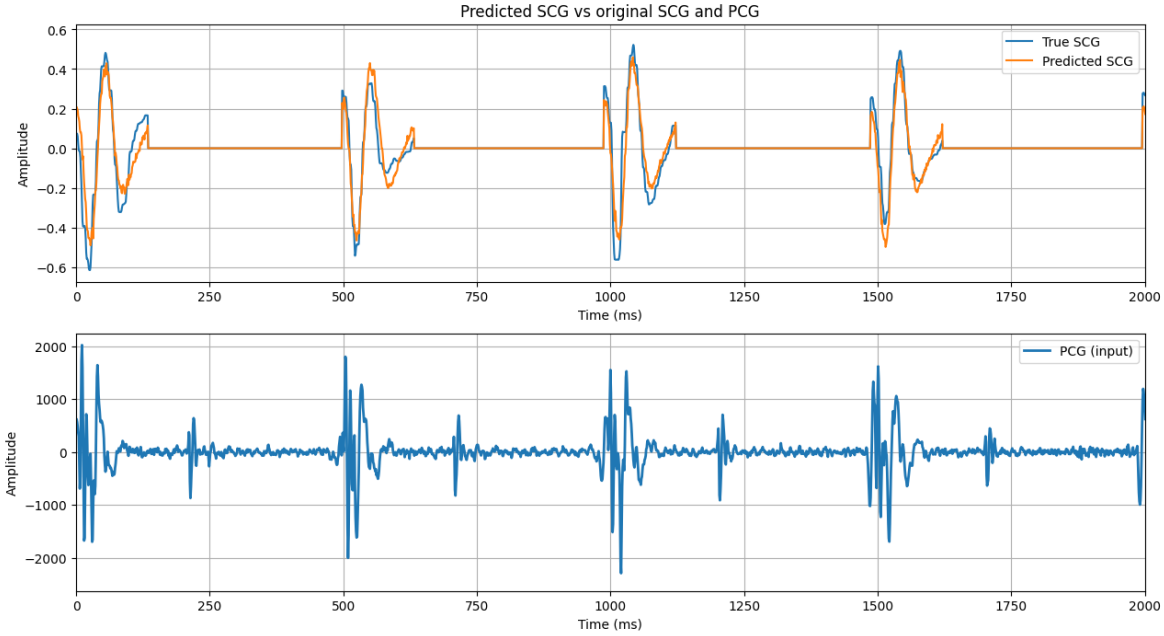

Figure 50. Results from Transformers, comparison between original and predicted SCG (top) with corresponding PCG (bottom)

The loss in this model decreases throughout all the epochs, which indicates that the model keeps learning new information until the end (Figure 51). There is a more significant drop in loss during the first 25 epochs, after which its slowly starts decreasing until almost 0.

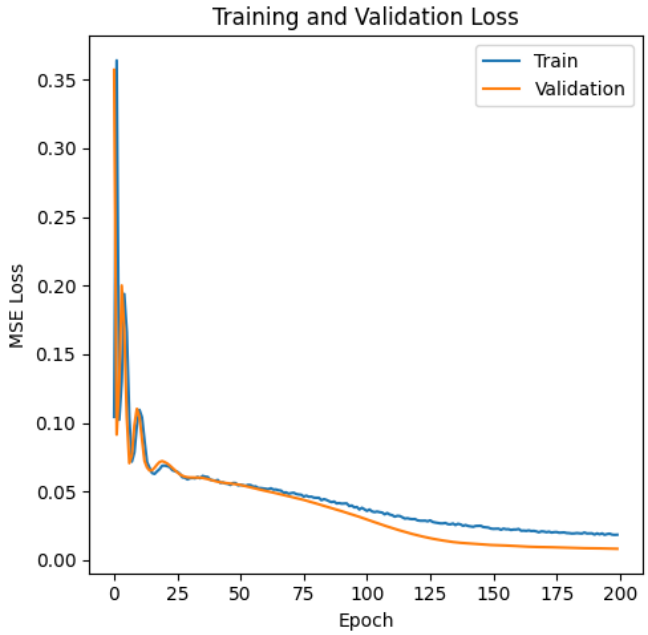

Figure 51. Results from Transformers, train and validation loss throughout the epochs

The R^2 score also has a significant increase, as shown in Figure 52. Although it starts more negative than any of the previous models, it quickly jumps to positive values and then keeps improving to almost 1. This high R^2 score is reflected in how well the predictions of this model are.

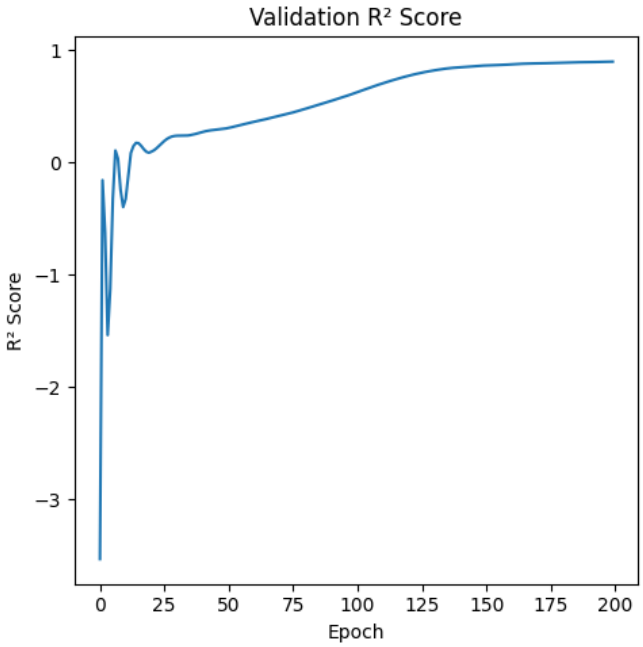

Figure 52. Results from Transformers, validation R^2 score throughout the epochs

The transformer model shows the best performance among the three architectures. These results indicate that transformers are the most promising for this signal mapping problem and is the most appropriate model for the next step of using PCG data from headphones.

6.4.4 IN-EAR HEADPHONE DATA MODEL

Using the phone data shows very promising results but it is not the most comfortable way to record these signals. To encourage patients to monitor their heart rate frequently, the data collection method must be simpler. For this reason, data was recorded using the in-ear microphone in headphones. These are able to record PCG data which can be mapped to SCG using the transformers model previously presented

The first step is to choose the best PCG signal and filter it to remove noise, as shown in Figure 53.

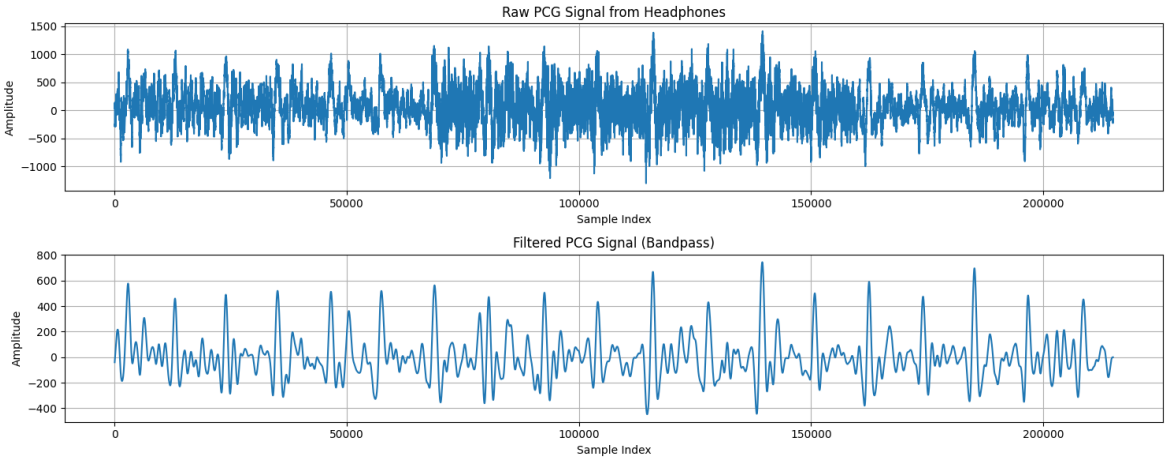

Figure 53. Headphone PCG data before and after filtering

Next, the signal had to be cropped to only the individual heart beats. This was done similarly to the previous data but detecting peaks (Figure 54) and cropping the signal around them.

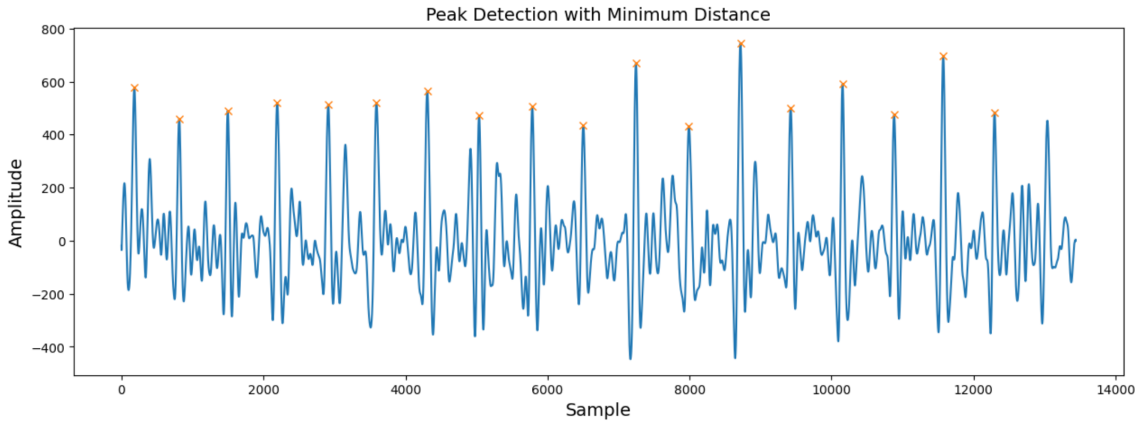

Figure 54. Peak detection for segmentation in headphone data

While the reconstructed signals may appear visually similar across samples, there is noticeable variability in the amplitude and morphology of the individual heartbeats. Due to the lack of corresponding ground truth SCG data for these segments, quantitative performance metrics could not be computed. However, based on the model's behavior during training and the qualitative assessment of the predicted waveforms, the transformer appears

to generalize reasonably well to unseen data, capturing the global structure of the SCG despite some variation in finer details.

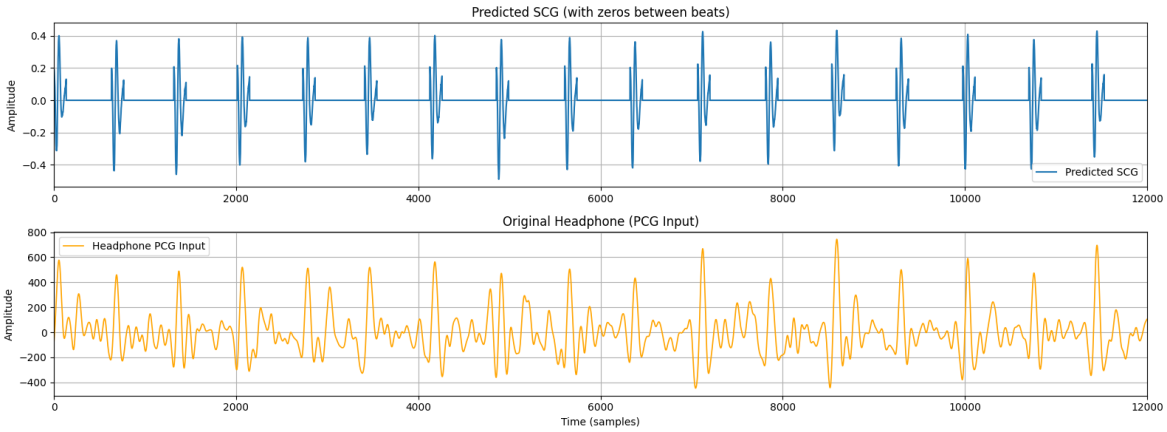

Figure 55. Headphone SCG signal prediction with transformer model

This latest experiment demonstrates the feasibility to use headphones to obtain a PCG signal of the heart in an easy and convenient way, and then use the model developed in this section (based on transformers) to obtain the corresponding SCG signal. Doctors will be able to analyze and produce a heart diagnosis out the SCG signal.

Chapter 7. CONCLUSIONS AND FUTURE WORK

In conclusion, the main objectives of the project, which included the development of an active learning model to predict failures in the portable ECMO machine, were completed. To this end, the model was trained with **in vitro data obtained from water loops** installed in the lab. It was possible to collect data of normal operations and to simulate different failure conditions. A comparison study of different algorithms was carried out.

One the approach of active learning was accepted, the model was subsequently validated using **in vivo data from sheep experiments**, that also generated normal condition and anomaly data in physiologically realistic scenarios.

As a complementary objective, the ability of the system to estimate in real time the respiratory and heart rate from signals recorded during testing was investigated. **Breathing rate** can be obtained through deep analysis of the same pressure signal used to monitor the ECMO device. It wasn't necessary to install new sensor or additional measuring equipment to obtain the respiratory information.

For **heart health monitoring**, reconstruction models capable of inferring key physiological variables such as the Seismocardiogram SCG from the Phonocardiogram PCG, in a completely non-invasive manner and without requiring additional effort on the part of the patient, were proposed. SCG signal can be analyzed by doctors to diagnose the heart, while the PCG signal can be easily collected with standard headphones.

The results obtained, both in failure prediction and vital signs estimation, support the potential of the proposed approach to improve patient monitoring and increase system safety. However, the accuracy, robustness and generalization of the models still need to be optimized before clinical implementation can be considered.

As lines of **future work**, it is proposed to finalize and publish the study focused on cardiac rhythm reconstruction, with the aim of consolidating the results obtained and providing additional evidence on the viability of the approach. In addition, the integration of a greater number of sensors in the system is contemplated to increase sensitivity and improve the detection of relevant physiological events. Additional testing in more diverse clinical or preclinical settings will also be necessary to validate the performance of the system under more realistic conditions and with greater inter-individual variability. Finally, we plan to explore the regulatory feasibility of the device, initiating the necessary procedures for its possible approval by entities such as the FDA, with a view to its future application in real clinical contexts.

Overall, this work lays the foundation for the development of an intelligent monitoring and failure detection system for portable ECMO. Although there are still steps to be completed, the progress achieved demonstrates the potential of the integration between machine learning and physiological sensors to improve patient safety and monitoring in critical environments.

Chapter 8. BIBLIOGRAPHY

Abrams, D., Madahar, P., Eckhardt, C. M., Short, B., Yip, N. H., Parekh, M., Serra, A., Dubois, R. L., Saleem, D., Agerstrand, C., Scala, P., Benvenuto, L., Arcasoy, S. M., Sonett, J. R., Takeda, K., Meier, A., Beck, J., Ryan, P., Fan, E., … Trindade, A. (2022). Early Mobilization during Extracorporeal Membrane Oxygenation for Cardiopulmonary Failure in Adults Factors Associated with Intensity of Treatment. *Annals of the American Thoracic Society*, *19*(1), 90–98. https://doi.org/10.1513/ANNALSATS.202102-151OC

Abrams, D., Schmidt, M., Pham, T., Beitler, J. R., Fan, E., Goligher, E. C., McNamee, J. J., Patroniti, N., Wilcox, M. E., Combes, A., Ferguson, N. D., McAuley, D. F., Pesenti, A., Quintel, M., Fraser, J., Hodgson, C. L., Hough, C. L., Mercat, A., Mueller, T., … Brodie, D. (2020). Mechanical ventilation for acute respiratory distress syndrome during extracorporeal life support research and practice. *American Journal of Respiratory and Critical Care Medicine*, *201*(5), 514–525. https://doi.org/10.1164/RCCM.201907-1283CI

Arias, F., Zambrano Nunez, M., Guerra-Adames, A., Tejedor-Flores, N., & Vargas-Lombardo, M. (2022). Sentiment Analysis of Public Social Media as a Tool for Health-Related Topics. *IEEE Access*, *10*, 74850–74872. https://doi.org/10.1109/ACCESS.2022.3187406

Bhattacharjee, P., & Mitra, P. (2021). A survey of density based clustering algorithms. *Frontiers of Computer Science*, *15*(1), 1–27. https://doi.org/10.1007/S11704-019-9059-3

Cacciarelli, D., Kulahci, M., Gama Davide Cacciarelli, J. B., & Kulahci muku, M. (2023). Active learning for data streams: a survey. *Machine Learning 2023 113:1*, *113*(1), 185–239. https://doi.org/10.1007/S10994-023-06454-2

Choudhary, T., Bhuyan, M. K., & Sharma, L. N. (2019). Orthogonal subspace projection based framework to extract heart cycles from SCG signal. *Biomedical Signal Processing and Control*, *50*, 45–51. https://doi.org/10.1016/J.BSPC.2019.01.005

Dutta, K., Lenka, R., Nayak, S. R., Khandual, A., & Bhoi, A. K. (2021). MED-NET: A novel approach to ECG anomaly detection using LSTM auto-encoders. *International Journal of Computer Applications in Technology*, *65*(4), 343–357. https://doi.org/10.1504/IJCAT.2021.117277

Hayes, K., Holland, A. E., Pellegrino, V. A., Mathur, S., & Hodgson, C. L. (2018). Acute skeletal muscle wasting and relation to physical function in patients requiring extracorporeal membrane oxygenation (ECMO). *Journal of Critical Care*, *48*, 1–8. https://doi.org/10.1016/J.JCRC.2018.08.002

Kottke, D., Krempl, G., & Spiliopoulou, M. (2015). Probabilistic Active Learning in Datastreams. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9385*, 145–157. https://doi.org/10.1007/978-3-319-24465-5_13

Lahiri, B., & Tirthapura, S. (2018). Stream Sampling. *Encyclopedia of Database Systems*, 3782–3787. https://doi.org/10.1007/978-1-4614-8265-9_372

Liu, S., Xue, S., Wu, J., Zhou, C., Yang, J., Li, Z., & Cao, J. (2023). Online Active Learning for Drifting Data Streams. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(1), 186–200. https://doi.org/10.1109/TNNLS.2021.3091681

Mayabadi, S., & Saadatfar, H. (2022). Two density-based sampling approaches for imbalanced and overlapping data. *Knowledge-Based Systems*, *241*, 108217. https://doi.org/10.1016/J.KNOSYS.2022.108217

Miller, E. M. (2020). Using Continuous Glucose Monitoring in Clinical Practice. *Clinical Diabetes : A Publication of the American Diabetes Association*, *38*(5), 429. https://doi.org/10.2337/CD20-0043

Nabih-Ali, M., El-Dahshan, E.-S. A., & Yahia, A. S. (2017). Heart Diseases Diagnosis Using Intelligent Algorithm Based on PCG Signal Analysis. *Circuits and Systems*, *08*(07), 184–190. https://doi.org/10.4236/CS.2017.87012

Nielsen, M. A. (2015). *Neural Networks and Deep Learning*. Determination Press. http://neuralnetworksanddeeplearning.com

Settles, B. (2009). *Active Learning Literature Survey*. https://minds.wisconsin.edu/handle/1793/60660

Sibai, R. El, Chabchoub, Y., Demerjian, J., Kazi-Aoul, Z., & Barbar, K. (2016). Sampling algorithms in data stream environments. *Proceedings - 2016 International Conference on Digital Economy: Emerging Technologies and Business Innovation, ICDEc 2016*, 29–36. https://doi.org/10.1109/ICDEC.2016.7563142

Skotte, J. H., & Kristiansen, J. (2014). Heart rate variability analysis using robust period detection. *BioMedical Engineering Online*, *13*(1), 1–11. https://doi.org/10.1186/1475-925X-13-138/FIGURES/4

Valapour, M., Lehr, C. J., Schladt, D. P., Smith, J. M., Swanner, K., Weibel, C. J., Weiss, S., & Snyder, J. J. (2024). OPTN/SRTR 2022 Annual Data Report: Lung. *American Journal of Transplantation*, *24*(2), S394–S456. https://doi.org/10.1016/J.AJT.2024.01.017

van Diepen, A., Bakkes, T. H. G. F., De Bie, A. J. R., Turco, S., Bouwman, R. A., Woerlee, P. H., & Mischi, M. (2021). A Model-Based Approach to Synthetic Data Set Generation for Patient-Ventilator Waveforms for Machine Learning and

Educational Use. *Journal of Clinical Monitoring and Computing*, *36*(6), 1739–1752. https://doi.org/10.1007/s10877-022-00822-4

Vaswani, A., Brain, G., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems*, *30*.

ANEXO I: PROJECT ALIGNMENT WITH THE SUSTAINABLE DEVELOPMENT GOALS

This research aligns with SDG 3 (Good Health and Well-being) by enhancing the reliability of ECMO systems, which are crucial for treating patients with severe cardiac and respiratory failure. Improved failure detection mechanisms contribute to better patient outcomes by enabling timely interventions and reducing preventable complications. Moreover, such real-time monitoring systems may allow patients to reside at home instead of being hospitalized, which drastically improves their quality of life.

Additionally, the research aligns with SDG 9 (Industry, Innovation, and Infrastructure) by advancing medical technology through the integration of machine learning in life-support systems. A further consideration is the long-term impact of AI-driven healthcare innovations. AI-assisted ECMO monitoring can contribute to the development of smart hospitals, where predictive analytics improve patient care and operational efficiency. This aligns with the overarching goals of SDG 3 by not only improving immediate health outcomes but also strengthening healthcare systems globally.