Informatics Institute of Technology
In Collaboration With
University of Westminster, UK

University of Westminster, Coat of Arms

Health Status Assessment in Remote Patient Monitoring Systems using Hybrid Machine Learning

A Project Proposal by
Mr. Hammadh Yusuf Mohamed Arquil
w1761780 / 2018128

Supervised by
Ms. Divya Premanantha

October 2022

This project proposal is submitted in partial fulfilment of the requirements for
the BSc (Hons) Computer Science degree at
the University of Westminster.

CONTENTS

List of Tables

List of Figures

List of Acronyms

RPM: Remote Patient Monitoring

ML: Machine Learning

DNN: Deep Neural Network

RNN: Recurrent neural network

LSTM: Long short-term memory

OOAD: Object-oriented Analysis and Design

RMSE: Root Mean Squared Error

MIT: Massachusetts Institute of Technology

PHC: Portable Healthcare System

1 INTRODUCTION

The purpose of this proposal documentation is to provide the reader with an overview of the current state of remote patient monitoring solutions in healthcare delivery and proposes a hybrid machine learning model to increase the performance and help automate the health status risk assessment process which includes diagnosis and prognosis of a user, which otherwise would require the specific thresholds to be set manually by experts, also allowing the system to alert the relevant healthcare providers.

This document defines the problem, the gap in the available literature, its challenges, and the strategic plan that the author would follow over the next several months to present the deliverables.

2 PROBLEM BACKGROUND

As treatment and monitoring of diseases require people to go frequently to hospitals, the healthcare industry dramatically changes and constantly advances towards a more technological transition, remote monitoring and remote healthcare are rapidly growing fields of study and research (Malasinghe, Ramzan and Dahal, 2019). Remote healthcare which could be divided into several domains such as telehealth, mobile health, etc. refers to the use of techniques and technologies to monitor patients situated outside the hospital environment. Remote Patient Monitoring allows healthcare providers and caretakers to view and monitor patient health information remotely from a distance and allow them to take the necessary steps to prevent the patient from deteriorating (El-Rashidy et al., 2021). The usage of these monitoring systems contributes to reduced healthcare costs by further decreasing the number of hospital inpatient admissions and usage of facilities (Dhinakaran et al., 2022). Also, with the usage of the remote patient monitoring solutions, these resources can further improve the quality of life of its users (Salehi et al., 2020).

In the current age of healthcare delivery, there is a huge importance and need for analyzing and assessing medical data with better performance and greater accuracy of medical assessments by providing diagnoses and predictions done with the usage of machine learning models especially when humans cannot directly recognize certain abnormalities (Jayatilake and Ganegoda, 2021). Assessing a patient's health status ahead of time would heavily improve a patient's chances of recovery as a late diagnosis could lead to several serious complications on the patients' health (Nallakaruppan and Kumaran, 2020). It is therefore clear that the neglect of accurate or late diagnosis often places the patient at the risk of poor consequences, so early detection of such deterioration enables immediate involvement from medical personnel (Elliott and Endacott, 2022).

3 PROBLEM DEFINITION

Traditional healthcare delivery systems require doctors to view the patient, diagnose the disease, and advise the future course of action for the patient (Vinutha, Kavyashree and Raju, 2022) although these methods pose some difficulty in terms of patient's mobility and daily workflow (Malasinghe, Ramzan and Dahal, 2019). Traditional remote patient monitoring further assists in that data transaction in healthcare delivery allowing doctors to assess the data and make decisions for the patient from afar while allowing patients to continue with their daily workflow. These systems too have their own limitations, one of them being due its remote nature certain emergencies cannot be accurately predicted or diagnosed prior. With the usage of machine learning algorithms, this limitation can be resolved to allow us to automate this process of analyzing the data by detecting abnormalities in the patient data early on in time and alerting the relevant healthcare providers of the patient's condition (Lata Sahu et al., 2021).

Due to the nature of the data, available machine learning model integrations with RPM systems point out a number of performance concerns. (Tabassum et al., 2020) proposed a hybrid encoding technique which they used in conjunction with the Random Forest algorithm, although they had received sufficient accuracy, they suffered from overall deficient performance in their solution for large datasets due to the large covariance between the techniques used. (Gontarska et al., 2021) developed a deep neural network model to calculate risk scores for cardiovascular diseases using vital signs collected however they also suffered from mediocre performance due to the time series nature of the data used. Having said that, we understand that one of the root issues that most of the available solutions for integrating machine learning into RPM suffer from the performance of machine learning models leading to a lowered efficiency of such systems.

3.1 Problem Statement

Existing solutions for assessments utilizing machine learning models embedded into remote patient monitoring systems have a series of model performance difficulties resulting in higher false positives and poor efficiency of such systems.

4 RESEARCH MOTIVATION

Recent advances in machine learning and data analysis have shown promising results in providing diagnosis and having the ability to predict specific diseases ahead of time, although most of them still have performance issues when it comes to the nature of the datasets. Since there is a need for cheaper healthcare, especially in third-world countries that cannot afford to use up their resources during a health crisis, providing assessments without using up hospital resources helps contribute to the domain. The author had worked with remote patient monitoring (RPM) systems and solutions prior to this proposal which has motivated this research to improve workflow efficiency and reduce healthcare costs using healthcare informatics.

5 RELATED WORK

Table 1: Related Work

Citation	Summary	Technique Used	Limitations	Contribution
(Lata Sahu et al., 2021)	Created a Cloud-based monitoring system that provides an alert notification when a vital sign hits either the upper or lower threshold set within the application settings	This study uses a Rule-based model to return an alert to the doctors.	The study requires care providers to manually set threshold parameters. It also does not take in the severity of vital sign data based on the patient's previous medical data.	Measured different physiological parameters using remote patient monitoring with real-time data transmission
(Tabassum et al., 2020)	Proposed an encoding technique to improve machine learning performance to	Used Random Forest with a novel encoding technique called One-hot frequency	In this study, the large covariance distance between the two techniques used in the developed hybrid	One-hot frequency Encoding is a hybrid technique between one-hot encoding and

	predict health status using data collected from a portable healthcare system (PHC)	encoding technique	technique causes overall poor performance.	frequency encoding. Used alongside with Random Forest algorithm, it had outperformed several other algorithms that did not use the hybrid encoding technique.
(Gontarska et al., 2021)	Created a deep neural network model to estimate risk scores based on vital parameters, trained on a dataset for cardiovascular diseases	Used a deep neural network (DNN)	The model developed is not patient-specific so its accuracy levels could be improved. Due to the time series nature of the dataset, the performance of the model can also be improved.	The study proposed a deep neural network model to estimate risk scores based on vitals and the model had received an AUROC score of 0.84
(Chang, Chang and Pourhomayoun, 2019)	Developed a critical vital sign risk prediction system using Recurrent Neural Network and Long Short-Term Memory to predict vital signs	Used Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM)	The performance of the system developed can be improved using Generative Adversarial Networks (GAN) to generate new data to build a	They proved that RNN-LSTM had provided the best accuracy results out of several classification algorithms in providing risk prediction.

	of ICU patients 1-hour in advance		semi-supervised model	
(Vinutha, Kavyashree and Raju, 2022)	Applied machine learning algorithms to a heart disease-related RPM system to classify a patient's health status as risk or non-risk while reducing false alarms by only alarming the healthcare providers when they observe repeated major abnormalities	The study used the K-Nearest Neighbor (KNN) technique	The results showed that the solution required a larger dataset from a population of people to become effective.	Classified patient health status and alerted only when repeated major abnormalities were detected to reduce the number of false positives. After an algorithm comparison, KNN had outperformed Naïve Bayes with an accuracy of 80% compared to 52%

6 RESEARCH GAP

Based on existing research, Table 1 specifies a need for an optimized and automated method of evaluating patients' health status with improved performance in remote patient monitoring systems leveraging machine learning models without relying on threshold limitations defined by doctors. (Lata Sahu et al., 2021). Due to the characteristics of the dataset used, some of the available studies have shown that their models did not give the desired performance. (Gontarska et al., 2021).

However, hybrid models in particular have not previously been extensively employed in remote patient monitoring systems; as a result, in this research, the author proposes a novel hybrid machine learning model that would diagnose the patient's vital signs to show the present status level, forecast the future condition of the patient's health that it is progressing towards, and utilize RPM systems to inform the healthcare personnel in advance to stop further patient deterioration.

7 RESEARCH CONTRIBUTION

7.1 Technological Contribution

This project proposes an implementation of a novel hybrid machine-learning model which will be developed to perform a health status risk assessment of a patient using vital signs collected in a remote patient monitoring system. Given the nature of the medical data employed, it is expected that this novel hybrid prediction model will perform more effectively than previous works that have offered similar outcomes.

7.2 Domain Contribution

The author proposes a comprehensive remote patient monitoring system capable of assessing the current health risk status based on patient vital signs, providing a prognosis of patient's condition, and alerting relevant healthcare providers to take appropriate action to prevent future deterioration.

8 RESEARCH CHALLENGE

The author faces a few obstacles during the project's timeline, some of which are listed below. Given the lack of literature currently available on health status risk assessment using hybrid machine learning models in remote patient monitoring systems, the following challenges are some of those faced by the author:

- Receiving access to a large critical healthcare dataset and undergoing required training to handle the data provided.
- Identifying the limitations of available literature to develop a novel framework that would provide an optimized performance during diagnosis and prognosis.
- Analyzing and utilizing appropriate tools, techniques, technologies and libraries to develop the hybrid machine learning model which will accept patient vital signs

9 RESEARCH QUESTIONS

- **RQ1**: What improvements in hybrid machine learning can be taken into consideration to maximize RPM performance?

- **RQ2**: What benefits can the adopted hybrid model bring to RPM systems?
- **RQ3**: How can the proposed system assist medical professionals in delivering better healthcare?

10 RESEARCH AIM

This research aims to conduct the design, development, and evaluation of a hybrid machine-learning model to be used for assessment using both diagnosis & prognosis of the health status of a patient using remote patient monitoring.

By utilizing the novel hybrid machine learning model that has been proposed, the research primarily intends to improve the performance and accuracy of health risk status assessments in remote patient monitoring systems.

11 RESEARCH OBJECTIVE

Table 2: Research Objectives

Objective	Description	Learning Outcome	Research Questions
Literature Review	Gather and review relevant information by reading, comprehending, and assessing previously published work. • **RO1**: Conduct investigations and studies done on existing remote patient monitoring systems. • **RO2**: Analyze the requirement of specialized prediction models within RPM systems • **RO3**: Obtain and review insights on the architecture of systems in remote patient monitoring	LO2, LO4, LO5	RQ1
Data Gathering and Analysis	Collect and analyze the requirements of the project using relevant tools and techniques available	LO1, LO2, LO3	RQ1, RQ2, RQ3

	• **RO1**: Gather requirements from the architecture of RPM systems. • **RO2**: Collect available data on patient medical data and vital signs • **RO3**: Get insights and feedback from available technology and domain experts		
Research Design	Create a system design that can effectively address the problems that have been identified. • **RO1**: Design an optimized hybrid framework capable of both diagnosis and prognosis for patient health status • **RO2**: Design a model with support for prognosis and diagnosis inputs.	LO1	RQ2, RQ3
Implementation	Implement the proposed system that can address the research gaps available • **RO1**: Implement a novel hybrid machine learning model capable of assessing patient health status efficiently with fewer false positives. • **RO2**: Implement a remote patient monitoring system and integrate the developed model.	LO1, LO5, LO6, LO7	RQ1, RQ2
Testing and Evaluation	Effectively evaluate the implemented model and the RPM system using recommended techniques. • **RO1**: Create a test plan with test cases, unit testing, performance testing and integration testing.	LO4	RQ2, RQ3

	• **RO2**: Test and evaluate the model against several benchmarking		
Documentation	Document the progression of the research project.	LO6, LO8	RQ1

12. PROJECT SCOPE

Based on the available timeframe for the research project, the scope is as follows:

12.1. In-scope

- Implementation of a hybrid machine-learning model capable of diagnosis and prognosis.
- Implementation of a remote patient monitoring system with a graphical user interface capable of taking patient input remotely.
- Integration of the proposed hybrid model with the RPM system to inform relevant healthcare providers of the patient's current health status and alerts when a prognosis is critical.
- Evaluation & comparison of the proposed system against existing solutions

12.2. Out-scope

- Integration of wireless IoT medical devices for automated input without relying on manual input from patients leading to improved accuracy of data collected.
- Ability to consider other additional external medical factors during prediction, such as previous health records, allergies, etc.
- Integration and support of the RPM system in the emergency department of hospital to monitor patients admitted.

12.3. Prototype Diagram

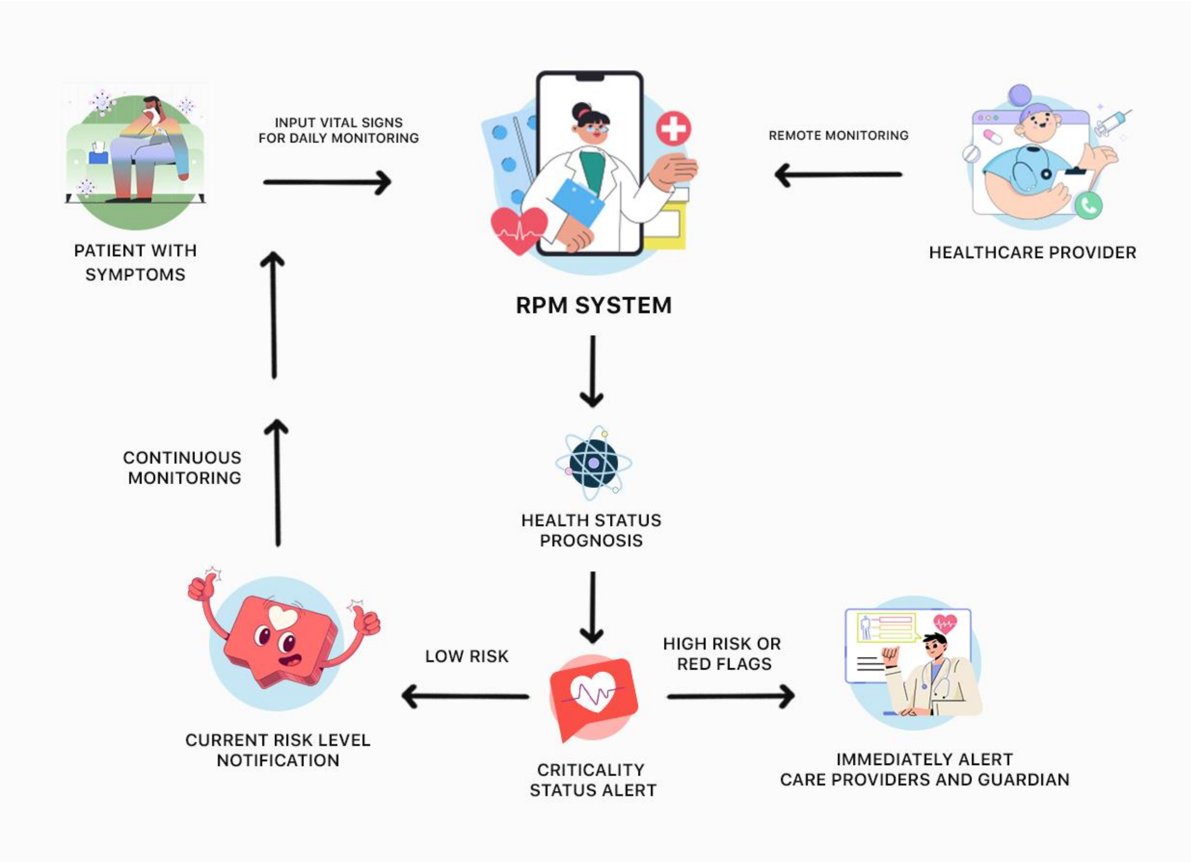

Figure 1: Prototype Workflow Diagram (Self-Composed)

13. METHODOLOGY

13.1. Research Methodology

The following research methodologies were evaluated and selected from a predefined Research Onion Model by (Saunders, Lewis, and Thornhill, 2007).

Table 3: Research Methodology

Research Philosophy	The research revolves around watching evaluation measures and feedback gathered through an iterative procedure, hence the author chose **Pragmatism** as the research philosophy.
Research Approach	The **Inductive** strategy has been chosen since it seeks to address the issue through trial and error.

Research Strategy	**Surveys, Interviews and Document Analysis** will be the primary research strategy to assist with evaluation, testing and feedback.
Research Choice	The **Mixed Method** would be used as it suits the proposed research for the required analysis
Time Horizon	Since the data will be collected all at once during the study analysis phase, the **Cross-sectional** time horizon was selected as the best choice.
Techniques and Procedures	The techniques used will be a form of **Data Collection and Analysis** through trial & error, analyzing similar solutions and literature.

13.2. Development Methodology

13.2.1. Life Cycle Model

Agile was chosen as the approach for the research and development life cycle because the project requires extensive iterative development.

13.2.2. Design Methodology

Object-Oriented Analysis and Design (OOAD) was chosen as the design methodology for this project as it is known for its support of reusability and increments to the application.

13.2.3. Evaluation Methodology

According to the proposed research, the **Classification Accuracy** level and **Root Mean Square Error (RMSE)** are the best metrics for measuring the proposed model's performance.

13.3. Project Management Methodology

As the project requires numerous iterations for ongoing changes, the author has chosen to combine **Agile** and **PRINCE2** while also partitioning the work into multiple chunks to focus on each with a plan-based approach.

13.3.1. Gantt Chart

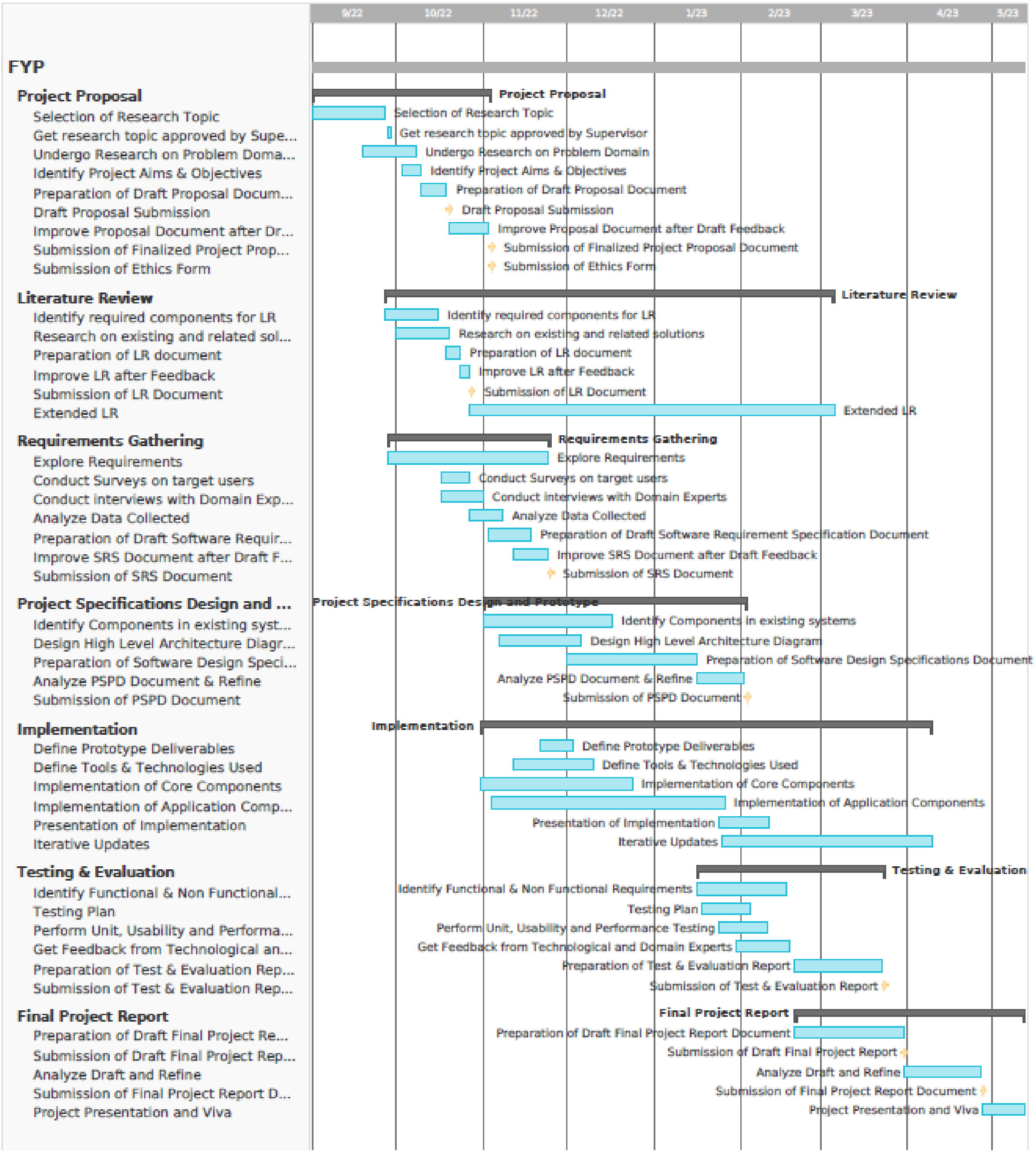

Figure 2: Gantt Chart

13.3.2. Deliverables

Table 4: Deliverables

Deliverable	Date
Literature Review Critical analysis of related work	27th October 2022
Project Proposal & Ethics Form	3rd November 2022

Initial document proposing the research to be conducted	
Software Requirement Specification (SRS) The guide that outlines the conditions that must be followed in order to develop the prototype and gather data	24th November 2022
Proof Of Concept & Implementation Presentation An initial Implementation of the proposed system.	23rd December 2022
Project Specifications Design & Prototype A prototype of the system with the core features and an attached document specifying the design followed & an overview of the implemented algorithm.	2nd February 2023
Test & Evaluation Report Documentation of test findings and evaluations conducted on the prototype.	23rd May 2023
Draft Project Report A draft submission of the Final Thesis to get evaluations.	30th May 2023
Final Thesis Final thesis submission with thorough project journey documentation.	27th April 2023

13.3.3. Resource Requirements

13.3.3.1. Hardware Requirements

- **Core i7 Processor (8th gen) or above** – To be able to handle models with intensive workloads
- **8GB RAM or above** – To be able to oversee multiple development environments, multitasking and model implementation
- **Disk storage of 10GB or more** – To store application code and data

13.3.3.2. Software Requirements

- **Operating System (Windows/macOS/Linux)** – Windows will be utilized as the default operating system because all necessary tools are readily available.
- **Python** – Will be used to create the proposed model, it is an all-purpose language with different tools and libraries designed specifically for data science
- **TensorFlow, Keras, and Scikit-learn** – Libraries to facilitate the data science modules in the model
- **Flutter** – To develop the client side of the application, Flutter is a development framework capable of cross-platform functionality leading to a single codebase
- **Firebase** – To facilitate the backend and storage module of the application with seamless integration to the client framework
- **Visual Studio Code, and Android Studio** – Will be used to facilitate the development of the application
- **Jupyter Notebook, and Google Colab** – Will be used as the development environment for building the prediction model.
- **Mendeley** – Used to manage references and research artifacts.
- **GitHub** – Will be used for Version Control during the implementation period of the proposed research.
- **Figma** – Will be used to create the high-fidelity design prototype of the application.
- **Microsoft Office, Draw.IO, Google Suite, and Notion** – Will be used in the documentation process to create reports, diagrams, figures, and backup files.

13.3.3.3. Skills Requirements

- Academic research and writing skills
- Work on the creation and implementation of a hybrid machine-learning model
- To be able to build scalable and optimal machine learning and deep learning models
- Being able to create client-side software with dynamically updating user interfaces based on expected outcomes

13.3.3.4. Data Requirements

- **Patient Vital Signs Dataset** – MIMIC IV ED Dataset (Johnson et al., 2022) provided by the Massachusetts Institute of Technology - requires credentialed access to the

PhysioNet platform and undergo the Data or Specimens Only Research training course certification for human subjects' research.

13.3.4. Risk Management

The following table identifies and highlights the possible risks that could occur during the project timeframe and how the author plans to mitigate them.

Table 5: Risk Mitigation Plan

Risk Item	Severity	Frequency	Mitigation Plan
Lose access to the code	5	2	The usage of Version Control solutions, and Cloud-based backups.
Corrupted Documentation	4	4	Backup and store all necessary files and documents on the cloud and external storage
Inability to deliver expected deliverables	4	2	Plan ahead and adhere to a list of priorities to deliver in accordance
Invalid hypothesis	3	2	Continue with the research as there will still be a contribution regardless
Lack of knowledge during the development of ML models	6	4	Learn from industry experts and gain insights

14 REFERENCES

Johnson, A., Bulgarelli, L., Pollard, T., Horng, S., Celi, L. A., and Mark, R. (2022) 'MIMIC-IV' (version 2.0), PhysioNet. Available at: https://doi.org/10.13026/7vcr-e114.

Saunders, M., Lewis, P. and Thornhill, A. (2003), ‘Research methods for business students, Essex: Prentice Hall: Financial Times. Available from https://toc.library.ethz.ch/objects/pdf_ead50/4/E57_7072090_TB-Index_005013522.pdf

Chang, Daniel, Chang, David and Pourhomayoun, M. (2019). Risk prediction of critical vital signs for ICU patients using recurrent neural network. *Proceedings - 6th Annual Conference on Computational Science and Computational Intelligence, CSCI 2019*. 1 December 2019. Institute of Electrical and Electronics Engineers Inc., 1003–1006. Available from https://doi.org/10.1109/CSCI49370.2019.00191.

Dhinakaran, M. et al. (2022). A System of Remote Patients’ Monitoring and Alerting Using the Machine Learning Technique. Available from https://doi.org/10.1155/2022/6274092

Elliott, M. and Endacott, R. (2022). The clinical neglect of vital signs’ assessment: an emerging patient safety issue? *Contemporary nurse*. Available from https://doi.org/10.1080/10376178.2022.2109494

El-Rashidy, N. et al. (2021). Mobile health in remote patient monitoring for chronic diseases: Principles, trends, and challenges. *Diagnostics*, 11 (4). Available from https://doi.org/10.3390/diagnostics11040607.

Gontarska, K. et al. (2021). Predicting Medical Interventions from Vital Parameters: Towards a Decision Support System for Remote Patient Monitoring. Available from http://arxiv.org/abs/2104.10085.

Jayatilake, S.M.D.A.C. and Ganegoda, G.U. (2021). Involvement of Machine Learning Tools in Healthcare Decision Making. *Journal of Healthcare Engineering*, 2021. Available from https://doi.org/10.1155/2021/6679512.

Lata Sahu, M. et al. (2021). Cloud-Based Remote Patient Monitoring System with Abnormality Detection and Alert Notification. *Mobile Networks and Applications*, 1, 16. Available from https://doi.org/10.1007/s11036-022-01960-4 [Accessed 4 October 2022].

Malasinghe, L.P., Ramzan, N. and Dahal, K. (2019). Remote patient monitoring: a comprehensive study. *Journal of Ambient Intelligence and Humanized Computing*, 10 (1), 57–76. Available from https://doi.org/10.1007/S12652-017-0598-X/TABLES/6.

Nallakaruppan, M.K. and Kumaran, U.S. (2020). Hybrid machine learning model for healthcare monitoring systems. *International Journal of Internet Technology and Secured Transactions*, 10 (5), 538–551. Available from https://doi.org/10.1504/IJITST.2020.109532.

Salehi, S. et al. (2020). Assessment of remote patient monitoring (RPM) systems for patients with type 2 diabetes: a systematic review and meta-analysis. *Journal of Diabetes and Metabolic Disorders*, 19 (1), 115–127. Available from https://doi.org/10.1007/S40200-019-00482-3/TABLES/4.

Tabassum, S. et al. (2020). A data enhancement approach to improve machine learning performance for predicting health status using remote healthcare data. *2020 2nd International Conference on Advanced Information and Communication Technology, ICAICT 2020*, 308–312. Available from https://doi.org/10.1109/ICAICT51780.2020.9333506 [Accessed 19 September 2022].

Vinutha, D.C., Kavyashree and Raju, G.T. (2022). Machine Learning–Assisted Remote Patient Monitoring with Data Analytics. *Tele-Healthcare*, 1–26. Available from https://doi.org/10.1002/9781119841937.CH1 [Accessed 10 October 2022].