CAMPUS: MAURITIUS

SCHOOL OF SCIENCE
AND TECHNOLOGY

Middlesex University

CST3990

Undergraduate Individual Project

DRAFT: LITERATURE REVIEW AND FIRST STEPS

AI-DRIVEN APPROACH TO LIFESTYLE RECOMMENDATIONS FOR CHRONIC KIDNEY DISEASE MANAGEMENT

by **Najya Hana Munshi** [M00906633]

Supervised by Dr. Aditya Santokhee

22nd February 2025

Contents

List of Figures

Chapter 1: Project Proposal

1.1 Background of Study

Chronic diseases have always caused a sense of foreboding and why not, society interprets it as an exile to a lifetime of suffering before an impending death sentence. However, that is a misconception. While chronic diseases like diabetes and kidney disease can raise significant health challenges, they are not necessarily that deadly if intercepted early on and preventative measures taken (Gruenberg, 1977). According to the World Health Organisation, the year 2021 witnessed around 43 million of deaths attributed to chronic diseases which marks a rise from 2020. This rise is concerning given the tremendous healthcare advancements such as genetic profile treatments and telemedicine which should theoretically be reducing mortality rates. The data also reveals a troubling trend especially in low and middle-income nations where healthcare infrastructure is inadequate (WHO, 2024).

Nonetheless, the healthcare landscape in high-income nations, where systems are supposedly more advanced, is not any better. Countries like Canada and Sweden have waiting times that often go up to 6 months before a specialist appointment (Lewis, 2024). Attempts like a wait-time guarantee policy have been established to solve this issue (ICI, 2024) however, for someone diagnosed with a chronic illness, especially for one such as chronic kidney disease (CKD) time is of the essence.

CKD is the perpetual and progressive loss of kidney function over the span of years (Hopkinsmedicine.org, 2024). It has different progressive stages, with the mildest stages being the first two and stage 5 representing kidney failure. A survey conducted in the UK reported that when CKD is first discovered in patients, more than 60% have already progressed to stage 3 or above (Amy, 2024). Consequently, prolonged waiting periods without any intervention will only exacerbate the progression of the disease. In the months leading up to the appointment, should the patient incorporate specific habits then the progression of the condition could be slowed down, and its symptoms mitigated.

1.2 Problem Statement

The advent of technology in the healthcare landscape has led to the adoption of recommendation systems in different aspects of medical practice for example in pharmaceutical assistance, nutrition and lifestyle management. Despite their growing popularity, a recommender system based on CKD management is a minimally explored area.

It has been shown that through adherence to a specific diet (Hahn et al, 2018; Naber and Purohit, 2021), and low-impact exercises (Villanego et al., 2020), the progression of CKD can be lessened. However, most studies only delve into the predictive aspect of CKD (Halder et al., 2024; Pankaj et al., 2021; Saif et al., 2024), and overlook the lifestyle changes that could be incorporated along with the prediction to guide the individual in tackling his affliction and almost certainly prevent severe reinal impairment.

A simple search on Google Scholar using the keywords ”CKD Recommendation System” and ”Diabetes Recommendation System” leaves the user with 17,500 and 1,320,000 results respectively. This affirms that a significant portion of literature have investigated recommendation systems for diabetes management while investigations into similar systems for CKD remain rather limited. It is within this research gap that our study positions itself, aiming to explore unaddressed potential features of recommendation systems in CKD management.

Alternatively, some CKD-based systems combine disease prediction with drug recommendation (Davy Tawadrous et al., 2011; Patwardhan et al., 2009). It is a practice that may raise significant concerns. Patients may have an allergy to a certain drug, or the drug recommended can adversely interact with other tablets the person takes. One notable case is IBM’s Watson recommending incorrect cancer treatments (Ross, 2018). Hence, lifestyle recommendations are considered to be a safer approach as they focus on non-invasive interventions such as dietary changes and physical activity.

In the traditional disease prediction system, the parameters like symptoms or blood test results must be entered manually one by one (Kanda et al., 2023). This not only takes time and also necessitates careful cross-checking to ensure the validity of the data, particularly for CKD prediction which requires several variables to make a precise prediction. For users not really conversant with technology, this can act as a barrier and discourage them from using the system (Langote et al., 2024).

In recent years, Google PlayStore has seen the addition of many mobile applications providing tailored Wellness and Lifestyle changes. Based on the data collected from IoT wearables, the person would be given real-time recommendations. Nonetheless, these applications lack the integration of disease prediction and disease prevention (Vairale and Shukla, 2018).

1.3 Project Description

'CKDVitaGuide' is a doorstep CKD diagnostics and lifestyle guide provider. The web application is designed primarily for undiagnosed CKD patients who may not have immediate access to healthcare. However, it can also be useful for individuals who suspect they might have CKD or want to check without visiting a doctor for an initial consultation.

The list below constitutes of the proposed system functionalities:

- **User Profile Creation:** To access the system's features the user will first need to create a profile.
- **Diagnosis:** Users can upload a picture of their blood test results. Through image processing and optical character recognition (OCR), the system will extract the relevant diagnostic values from the attachment. The likelihood of the selected disease will be predicted using a machine learning algorithm.
- **Lifestyle Recommendations:** If the prediction is positive, users will receive tailored recommendations which can include nutrition, exercise, hydration and alcohol intake.
- **Educational Resources:** A page will be dedicated to information on the types of blood tests required for diagnosing CKD by the system if suspected.
  All information and recommendations provided by the system will be backed by academic literature.

1.4 Aim and Objectives

1.4.1 Aim

The aim of this project is to create a system able to diagnose CKD in individuals by analysing their blood tests results and hereby recommend personalised lifestyle changes to slow the progression of the disease.

1.4.2 Objectives

The list below pertains to the main deliverables to achieve the aim of this project:

- Literature Review
- System Analysis and Design
- Implementation of System
- Evaluation

1.5 Key Activities of Project

1.5.1 Literature Review

The methodology employed for the literature review began with devising a list of different variations of keywords related to a disease prediction and healthcare recommendations system. This included "recommendation/recommender systems", "chronic disease prediction", "machine learning" amongst others. Using these keywords, a thorough search was conducted to look for relevant academic papers, books and articles on academic databases like Google Scholar, IEEE Xplore, ScienceDirect and Springer. Since the academic papers sometimes engage in unfamiliar concepts, the process also involved further research into comprehending those concepts and techniques through the Internet.

1.5.2 System Analysis and Design

Before creating the web-application, a thorough design to illustrate the different use cases of the system must be undertaken. The system architecture and relevant technology stack will be outlined through means of diagrams to illustrate the workflows between the functionalities and relationships of the system's components.

Following that, wireframes will be created, and they will be used as skeleton for the frontend development of the website.

1.5.3 Implementation of System

The website implementation will be overseen by the Waterfall development methodology.

Implementing the system will start off with creating the frontend of the web-application in alignment with the wireframes. Once this phase is completed and the relevant web-stack technology is decided, the backend development will follow. This will involve the execution of the system's different functionalities like user account creation, scanning blood test results and constructing a tailored plan based on disease prediction. The feature for automating the extraction of the relevant blood test parameters from the attachment will employ machine-learning-based OCR techniques. For disease prediction, a high performing model will be trained on a CKD related dataset.

1.5.4 Evaluation

Evaluation of the system will be divided into 2 stages. One stage will involve functional testing of the web application to ensure that input validation, forms and buttons are working as intended. The behaviour of the API endpoints should also be validated. The other stage covers the evaluation of the performance of the machine learning algorithms for CKD prediction. Several evaluation metrics such as Confusion Matrix, Accuracy, Sensitivity, Precision, F1 Score amongst others will be studied and compared for each algorithm to determine which one has the better performance. Given the time constraint for the realisation of this project, it would unfeasible to hold user acceptance testing with subjects such as doctors or CKD patients to obtain feedback, instead, a group of random users can be selected for this part of the usability testing.

1.6 Resources Declaration

Personal Laptop: Used for planning and research, writing the dissertation, code development, and model training.

Project Management Tools
SmartSheet: For creating Gantt Chart and tracking deliverables.
Zotero: To organise and annotate academic papers and articles

System Design Tools
SmartDraw: For the wireframes of the webstie and design of system architecture.
UMLet: To draw relevant UML diagrams

Development Tools
Colaboratory(Free Tier): For training jobs.
Visual Studio Code: For code development.
GitHub: For code backup and version control

Other Relevant Tools

Kaggle(Free Tier): For dataset search.
LaTeX: For writing and structuring the report.

1.7 Project Planning

1.7.1 Milestones

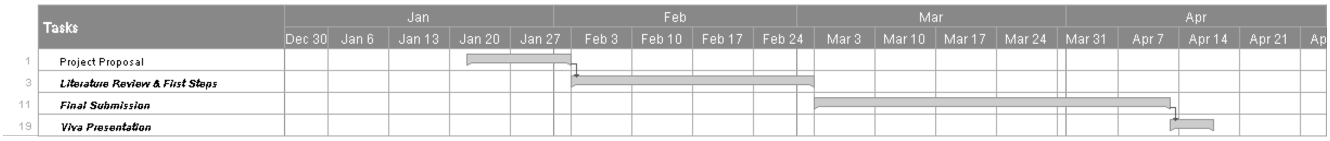

Figure 1.1: Milestones

1.7.2 Main Deliverables

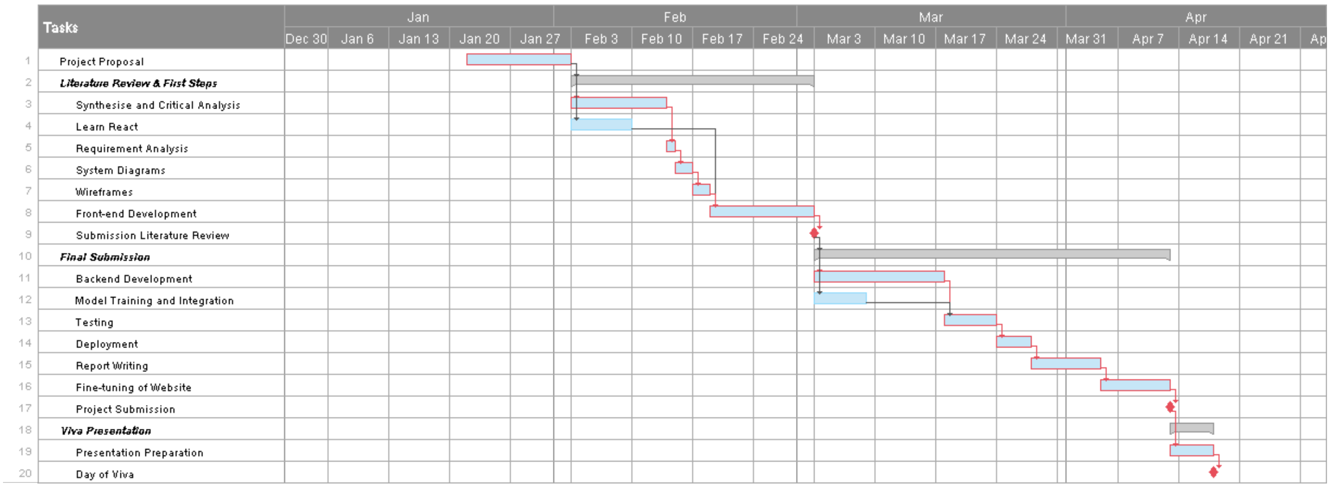

Figure 1.2: Gantt Chart

Chapter 2: Literature Review

The process of finding relevant sources for our literature review began with scrutinising reputable academic databases which included IEEE Xplore, ScienceDirect, Google Scholar and Springer. To obtain an analysis of domain-specific knowledge, a bibliographic review method was adopted. The search not only comprised keywords such as ”health recommendation systems”, ”chronic kidney disease”, ”prediction”, ”recommender systems”, ”lifestyle changes”, ”machine learning” but also extended to ”diet recommendations”, ”fitness recommendation”, ”drug recommendation”. Journals which were more than 5 years old have been eliminated. A total of 150 papers were retrieved and further refinement was done based on the quality of the journals by evaluating their abstract and ascertaining if they were well referenced with more than 15 sources. That left us with 44 papers which we glossed over the material to see if they could align with our research. Our final count was narrowed down to 14 papers that were deemed relevant to our research. These studies were listed in the *References* Section: 9 papers for CKD prediction, 2 papers for drug recommendation and 3 papers related to food recommendations.

2.1 Evaluation of Existing Research

2.1.1 System for CKD Progression and Mortality

Kanda et al. developed a web-based application for predicting CKD mortality. Users input patient data, such as estimated glomerular filtration rate (eGFR), and the system applies a machine learning model to assess the risk of end-stage CKD or death in diagnosed patients. To facilitate understanding of the model’s prediction, the system also categorises patients into high-risk and low-risk groups based on the Youden index.

The system sought to bridge the gap between research and clinical practice hence, the primary targets are healthcare providers.

Key Features and Strengths

1. Comprehensive Dataset and Validation

Unlike most CKD prediction studies that rely on the CKD dataset from UCI Machine Learning with around 400 samples for model training (Figueroa et al., 2024; Halder et al., 2024), Kanda et al. developed their prediction system using a larger dataset of approximately 3,500 samples. Their machine-learning models were trained on electronic medical records (EMR) from a Japanese medical school hospital. The dataset incorporates both first-visit data and time-series data which can help in tracking disease progression (Hirano et al., 2019).

Furthermore, the models were validated using a cohort study of approximately 27,000 CKD patients and demonstrated a high overall accuracy.

2. Patient Variable Selection

From the EMR, 22 variables were curated. These were carefully chosen with the aim that they are commonly measured factors in clinical use which will ensure the ease of implementation in a real-world setting (American Kidney Fund, 2021). From these 22 variables, the researchers varied the training of their models with smaller portions of the dataset but the performance of the model remained a solid one despite the limited data.

3. Compatibility with EMR

The authors claims to be the first study to test and implement AI models for CKD end-stage risk prediction. Their system address common barriers such as the integration of AI models with different hospitals' EMR systems (Luong, 2024). This could make the system's adoption in hospitals less complicated.

Limitations

1. No Lifestyle Consideration

While this research takes a deep dive into the risk of end-stage CKD, it completely overlooks the measures that can be taken to minimise the progression of the disease to fatality. Medical science has proven time and time again that incorporating lifestyle changes especially the dietary kind, can considerably prevent the patient's kidneys from reaching total failure and developing other comorbidities (Kelly et al., 2020; Neale et al., 2023).

2. System Restricted to Stage 5 Prediction

Risk prediction of stage 5 CKD in patients is important but the study could have been further comprehensive if predicting CKD in undiagnosed patients was also undertaken.

Too often, people have already reached moderate to severe loss of kidney functions when they are diagnosed (Amy, 2024).

3. Reported Data The system is designed to minimise manual data entry by automatically retrieving the variables from the patient's data on the EMR however, medical data records do not have a standardised format across healthcare services. Often, the data collected contains missing values and does not always come in a structured manner (Getzen et al., 2023). In the study undertaken, more than 20% of the missing values were excluded. This can have the potential to limit the accuracy of the model.

Relevance to the Problem

The research conducted by Kanda et al. highlights the importance of specific features like creatinine levels and eGFR to predict end-stage CKD in diagnosed patients. Our research will focus on diagnosing CKD in patients irrespective of their stage of progression but the specific features used in Kanda et al.'s model can be repurposed for our system. Kanda et al. also sought to limit manual data entry through means of linking their system to an institution's EMR, essentially making this process more efficient. Driven by the goal of enhancing efficiency, we will automate this process by implementing a system that scans and extracts relevant values from the user's blood test results.

2.1.2 Diet and Workout Recommendation System

The system is a android-mobile application that provides personalised workout and diet recommendations. With the widespread prevalence of diseases today, people have become more health-conscious. Rather than scouring the internet for dietary and fitness advice, which can often be overwhelming, this project aims to consolidate all those queries into a single platform. After calculating their Body Mass Index (BMI), the system will categorise the user into one of these 3 categories: overweight, underweight or healthy. Following that, it will tap into its machine learning models to present the user with a tailored diet and exercises.

The target users of this study are health-conscious individuals.

Key Features and Strengths

1. Personalised Recommendations
Based on user inputs such as age, height and weight, the system provides recommendations adapted to the person's needs. It will calculate the BMI and based on the

cluster assigned, suggest diet plans and workout videos. The system also considers user preferences, such as vegetarian and non-vegetarian which could make adopting a health-conscious lifestyle more attainable (Sandri et al., 2024).

2. Use of Machine Learning
The integration of K-means clustering and Random Forest classification allows the system to handle complex data and provide more reliable recommendations. Based on other studies, Random Forest appears to score high on the precision scale for lifestyle recommendations. This combination reduces overfitting and improves the reliability of the suggestions.

3. Healthy Living
The system promotes prevention is better than cure. It encourages users to adopt healthier lifestyles, addressing the rising cases of obesity and related diseases.

Limitations

1. Limited Scope of Data Input
The system relies on basic user input such as age, weight, height and fitness goals. These parameters do not capture the complexity that one's health condition can be in (Saye, 2024). Other important factors like a person's food allergies, glucose level, RBC and PCOS, can make all the changes proposed by this system irrelevant.

2. Evaluation of the Models
The authors of this project did not evaluate the performance of the K-Means clustering or the Random Forest model. Even when focusing on the application of these models for recommendation purposes, they neither justify their choice of algorithms nor provide any discussion on their validation. The efficacy of these models within the project's context remains unestablished.

3. Android Application
Another notable limitation is the system not being a cross-platform one. In a world where users often access applications on multiple devices such as laptops, tablets, this will essentially restrict its accessibility and reduce the number of users willing to engage with such a system (Claritee, 2025).

Relevance to the Problem

Both Sadhasivam et al.'s system and our proposed system will rely on the concept that an active lifestyle and the proper nutrition can help minimise the risk of diseases but where Sadhasivam et al. refer to diseases generally, we will be focusing on these lifestyle interventions to slow the progression of CKD. A 14-year clinical study has demonstrated that adhering to CKD-specific dietary models can reduce the risk of progression by 25% (Hu et al., 2021).

2.1.3 Study on Diagnosing CKD with ML

J. Figeuroa et al. came up with a refined solution for diagnosing and characterising CKD with machine learning in late 2024. The main purpose of the study was to investigate how feature selection regarding patient's symptomatic data can influence the model's accuracy. Unsupervised machine learning which is scarcely employed in CKD prediction was also delved into.

Their study fundamentally targets CKD patients and those at risk of developing the condition. Clinicians are another primary user group. It could assist them in making better prognoses.

Key Features and Strengths

1. Patient Clinical Characteristics
Patient clinical characteristics include specific symptoms such as swelling in the legs and comorbidities that are relevant to the diagnosis of CKD. At an early stage of 1 or 2, many patients are asymptomatic. Their blood test results do not reveal the disease and that could explain why most persons are diagnosed at a point when their kidney function is already severely compromised. Introducing these characteristics to the model will enhance the predictive model as can be evidenced by the 99.4-100% accuracy in J. Figeuroa et al.'s research.

2. High Performing Models
Variations in the configuration of Light Gradient Model presented the researchers with a model whose accuracy fluctuates very good between 99.4 to 100%. The model can be trusted in accurately identifying the disease in patients.

3. Validation Methods
Comprehensive methods were employed to evaluate the model. Many studies restrict their evaluation to one method either hold-out or cross-validation. Hold-out is the splitting of the dataset into a 'train' and 'test' category. It is the conventional 80-20 splitting used in machine learning model training. Cross-validation randomly splits the dataset into 'k' groups. For instance, a 5-fold cross validation is separated into 5 groups. The model is then trained and tested on 5 different occasions. Ultimately, each group becomes the test set at one point.

J.Figeuroa et al. applied both methods. This provides a more accurate assessment of the model's performance.

Limitations

1. Dataset Restrictions

The model utilised for CKD predictions was not trained on a substantial amount of data. Machine learning algorithms normally yield better model performance when confronted to real-world data with larger datasets. The researchers of this study used a dataset comprising only 400 samples, with 250 CKD cases and 150 non-CKD cases and this may additionally introduce imbalance bias.

2. Limited Comparison of Algorithms

The study noted the different high-performing algorithms used in CKD prediction but no attempt was made to train and evaluate the model using one or more of the machine learning algorithms. Light Gradient Model was selected and the model was trained accordingly. Comparisons could help in identifying more robust models.

3. No User Interface

While a user interface (UI) is not necessarily a limitation in the context of this study's objectives however, for clinical use in the real-world the development of a UI is crucial. The translation of the prediction results in a more accessible and interpretable way would ensure the predictive capibilities are effectively used by healthcare providers.

Relevance to the Problem

This study conforms to our efforts of achieving a highly accurate CKD predictive model. The findings in this paper present a robust model whose accuracy ranges between 99.4% to 100%. This will be significant in our research for considering the best-performing algorithms to predict CKD.

Our goal is to tackle CKD at its earliest stage and this aligns with J. Figeuroa et al.'s model.

2.2 Comparative Analysis

Features	**System for CKD Progression and Mortality**	**Diet and Workout Recommendation System**	**Study on Diagnosing CKD with ML**	**CKDVitaGuide**
User Interface	Web-based	Android	None	Web-based
Dataset Size	3500 samples	Undisclosed	400 samples	1600 samples
Disease Prediction	End-stage CKD in diagnosed patients	Health Status based on BMI	CKD diagnosis	CKD diagnosis
Input Data	Patient data from EMR	Self-report user data	No Interface	Blood Tests Results Image
Lifestyle Recommendations	None	Diet and Fitness Recommendations	None	Lifestyle Recommendations
Target Audience	Healthcare Providers	Health-conscious persons	CKD patients and clinicians	CKD patients
Blood Test parameters	Considered	None	Considered	Considered
Diet preferences	None	Vegetarian and Non-vegetarian options	None	None
Image Analysis	None	None	None	Yes
User Dietary Intake	None	Considered	None	None
Userability	Good	Good	None	Good

Table 2.1: Comparisons of Solutions

NOTE: Not yet completed.

2.3 Critical Analysis

The literature in this field has many studies carried on CKD prediction. However, most of these studies were based upon a CKD dataset provided by UCI Machine Learning Repository. This dataset contains a limited number of 400 samples and for training machine learning models, a large sample size is key in obtaining a model that generalises properly to unseen data. This is because with more data, more complex models are created (Krishnakumar, 2024). Another issue with this dataset is that it dates from 10 years ago now. Few studies have attempted to predict CKD occurrence using other datasets. Krishnamurthy et al. used Taiwan's National Health Insurance Research Database to implement their CKD predictive model. The limitation with their data samples is the imbalance bias, with 72, 000 non-CKD diagnoses samples against a meagre 18, 000 for CKD diagnoses. This can result in poor performance when identifying the minority class, which is the primary focus of this study. A more recent study utilised patients samples from an Ethiopian Hospital (Debal and Sitote, 2022). Since the binary-class distribution was yet again imbalanced with 16% for non-CKD, they adopted an oversampling resampling technique to balance the samples in the minority class with the majority class. Afterwards, the size of the dataset amounted to 2888. As the UCI Machine Learning Repository dataset has been extensively studied in previous research, our project will instead use a dataset with approximately 1,600 samples. However, a challenge anticipated is class imbalance and we plan to address this using Adaptive Synthetic Sampling.

For CKD prediction, supervised learning have gained more popularity than unsupervised learning amongst researchers. In cases where unsupervised machine learning techniques have been used, they are almost always k-means or KNN algorithms. Upon reviewing these studies, it can be noticed that these algorithms achieve high accuracy, often exceeding 0.9 (Antony et al., 2021; Figueroa et al., 2024). However, their models were trained on small datasets. KNN tends to perform well with smaller datasets, while k-means is sensitive to outliers. Given the nature of the dataset used, this explains why these unsupervised models appeared to generalise well. In reality, CKD is a complex disease with multiple stages of progression, and medical data often contains anomalies. The healthcare industry relies on transparency and trustworthiness, so prediction results must be clear and interpretable. For example, KNN makes it difficult to understand how predictions are made since it does not define clear decision boundaries or indicate which features are most important in classification.

For this reason, our project will consider only supervised learning techniques. Many researchers have employed differing techniques to obtain models whose performance are above 0.9. Some researchers have used the same UCI dataset but achieved poor model performance, while others using the same dataset obtained models with accuracy ranging from 0.95 to 0.99. This suggests that factors such as feature selection and validation approaches play a significant role in determining a model's performance.

NOTE: *To be continued. Will go more into details on feature selection and algorithm* *validation. Another theme will cover the automatic imputation of parameters from the blood test image to run the predictive model*

2.4 Summary of Findings

2.5 Proposed Solution

The proposed application, CKDVitaGuide, is a predictive and lifestyle recommendation system for chronic kidney disease. Given that CKD exhibits minimal symptoms in its early stages, the platform aims to provide an accessible preliminary assessment for individuals with limited or delayed access to healthcare. By offering timely insights, CKDVitaGuide seeks to mitigate the risk of disease progression for those who may otherwise face barriers to immediate medical consultation.

The application will be a web-based platform, any device with a browser and an Internet connection will be able to access it. Many studies have focused on CKD prediction but they lack an accompanying user interface. Our project will aim to fill the lack of user-friendly interface to enhance accessibility and usability.

There will be a page on the website which will cover the different research-backed blood tests required for our system to make a good diagnosis. Previously, no solution provides a clear overview for the user to understand why these specific blood tests are requested.

Our application will also allow users to bypass the repetitive and tedious process of manually reporting blood test parameters one by one. Instead, the system will scan the blood tests results from an attachment, extract the relevant parameters and automatically feed them for prediction. Current state-of-the-art research have yet to address this manual process.

CKDVitaGuide does not only predict the risk of CKD in a patient but it also accompanies the diagnosis with tailored dietary and fitness advice. Previously, most CKD prediction studies did not consider this aspect.

Chapter 3: First Steps

3.1 Development Methodology

The undertaking of this project has multiple layers such requirements gathering, system design, implementation of the web application and evaluation. Once requirements gathering has been done through research, the requirements of the project are unlikely to be subject to change and the remaining phases will follow sequentially. For the linearity of this project's life cycle, the Waterfall development methodology will be adopted. Given the 12-week timeframe for this project, the rigid and sequential approach of the Waterfall model is the most suitable choice to ensure timely completion within the strict deadline. It has also been reported that software projects falling under the healthcare industry tend to lean more towards the Waterfall methodology owing to the detailed documentation and standard industry compliance required (Groove Technology, 2025).

3.2 Functional and Non-Functional Requirements

3.2.1 Functional Requirements

Functional Requirements	Evidence
Create an account on the platform	When healthcare platforms are collecting patient data that includes protected health information (PHI), industry standards such as HIPAA mandates that they protect patient data, which can be achieved through user authentication systems (Office, 2022)
Log in to access user profile and system's features	This provides users with a centralised platform to view their diagnosis and personalised treatment plan (Microsoft, 2018). This can reduce redundant steps that requires the user run the diagnosis again in case they forgot about the specific changes recommended.
Start CKD diagnosis using blood test results: - **Receive positive results and lifestyle recommendations** - **Receive negative results**	Most CKD prediction studies overlook the lifestyle changes that can prevent the exacerbation of the condition. Our system will combine prediction with a personalised lifestyle intervention plan because medical research has proven that nutrition and physical activity can help minimise the damage to kidney functions.
Read about the blood tests needed for the system's diagnosis on the website	Clinical patient data selection is crucial to obtain a model that is not learning from irrelevant features for CKD diagnosis (Debal and Sitote, 2022). For this reason, the key variables most relevant for identifying CKD have been selected and this page will provide users with information on the specific blood test parameters required by the system and benefit from knowing why the tests are required.

Table 3.1: Functional Requirements

3.2.2 Non-Functional Requirements

The non-functional requirements of the system are as follows:

- CKDVitaGuide should support multiple users accessing and using its features simultaneously.
- The system should clearly explain the diagnostic method and the reasoning behind each blood tests variables request.
- The platform should have an intuitive user interface that makes the process of diagnosis as efficient as possible.
- The user's profile should be secured and resistant to brute force attacks.
- In case of system shutdown, no data should be lost.

3.3 UML Diagram

3.4 Tools and Technologies

3.5 Remaining Objectives

Chapter 4: References

American Kidney Fund. (2021). *Tests for kidney disease.* [online] Available at: https://www.kidneyfund.org/all-about-kidneys/tests-kidney-disease [Accessed 19 Feb. 2025].

Amy (2024). *Late diagnosis causes major harm to kidney patients.* [online] Kidney Research UK. Available at: https://www.kidneyresearchuk.org/2024/03/14/late-diagnosis-causes-irreversible-and-major-harm-to-kidney-patients/ [Accessed 1 Feb. 2025].

Antony, L., Azam, S., Ignatious, E., Quadir, R., Beeravolu, A.R., Jonkman, M. and De Boer, F. (2021). *A Comprehensive Unsupervised Framework for Chronic Kidney Disease Prediction.* [online] IEEE Access, 9, pp.126481–126501. Available at: https://doi.org/10.1109/access.2021.3109168 [Accessed 1 Feb. 2025].

Claritee (2025). *Ensuring a Seamless Cross-Platform Experience: Strategies and Best Practices - Team collaboration, UX & Functionality.* [online] Team collaboration, UX & Functionality - Blog by Claritee.io. Available at: https://claritee.io/blog/ensuring-a-seamless-cross-platform-experience-strategies-and-best-practices/?utm_source=chatgpt.com [Accessed 21 Feb. 2025].

Debal, D.A. and Sitote, T.M. (2022). *Chronic kidney disease prediction using machine learning techniques.* Journal of Big Data, 9(1), p.109. Available at: https://doi.org/10.1186/s40537-022-00657-5 [Accessed 19 Feb. 2025].

Tawadrous, D., Shariff, S.Z., Haynes, R.B., Iansavichus, A.V., Jain, A.K. and Garg, A.X. (2011). *Use of Clinical Decision Support Systems for Kidney-Related Drug Prescribing: A Systematic Review.* American Journal of Kidney Diseases, [online] 58(6), pp.903–914. Available at: https://doi.org/10.1053/j.ajkd.2011.07.022 [Accessed 22 Feb. 2025].

Figueroa, J. et al. (2024). *Diagnosing and Characterizing Chronic Kidney Disease with Machine Learning: The Value of Clinical Patient Characteristics as Evidenced from an Open Dataset.* Electronics, 13(21), p. 4326. Available at: https://doi.org/10.3390/electronics13214326 [Accessed 22 Feb. 2025].

Getzen, E., Ungar, L., Mowery, D., Jiang, X. and Long, Q. (2023). *Mining for equitable health: Assessing the impact of missing data in electronic health records.* Journal of Biomedical Informatics, [online] 139, pp.104269–104269. Available at: https://doi.org/10.1016/j.jbi.2022.104269 [Accessed 22 Feb. 2025].

Groove Technology (2025). *Waterfall Model In Software Projects: Best Practices.* [online] Available at: https://groovetechnology.com/blog/software-development/waterfall/ [Accessed 22 Feb. 2025].

Gruenberg, E.M. (1977) 'The failures of success', *The Milbank Memorial Fund Quarterly. Health and Society*, 55(1), p. 3. Available at: https://doi.org/10.2307/3349592.

Hahn, D., Hodson, E.M. and Fouque, D. (2018). *Low protein diets for non-diabetic adults with chronic kidney disease.* Cochrane Database of Systematic Reviews. [online] Available at: https://doi.org/10.1002/14651858.cd001892.pub4 [Accessed 22 Feb. 2025].

Halder, R.K. et al. (2024). *ML-CKDP: Machine learning-based chronic kidney disease prediction with smart web application.* Journal of Pathology Informatics, 15, p. 100371. [online] Available at: https://doi.org/10.1016/j.jpi.2024.100371 [Accessed 22 Feb. 2025].

Hirano, K., Kobayashi, D., Kohtani, N., Uemura, Y., Ohashi, Y., Komatsu, Y., Yanagita, M. and Hishida, A. (2019). *Optimal follow-up intervals for different stages of chronic kidney disease: a prospective observational study.* Clinical and Experimental Nephrology, 23(5), pp.613–620. [online] Available at: https://doi.org/10.1007/s10157-018-01684-4 [Accessed 22 Feb. 2025].

Hopkinsmedicine.org. (2024). *Chronic Kidney Disease.* [online] Available at: https://www.hopkinsmedicine.org/health/conditions-and-diseases/chronic-kidney-disease [Accessed 1 Feb. 2025].

Hu, E.A. et al. (2021). *Adherence to Healthy Dietary Patterns and Risk of CKD Progression and All-Cause Mortality: Findings From the CRIC (Chronic Renal Insufficiency Cohort) Study.* American Journal of Kidney Diseases, 77(2), pp.235–244. [online] Available at: https://doi.org/10.1053/j.ajkd.2020.04.019 [Accessed 22 Feb. 2025].

International Citizens Insurance (2024). *Sweden's healthcare system.* [online] Available at: https://www.internationalinsurance.com/health/systems/sweden.php#:~:text=Under%20Swedish%20law%2C%20patients%20should,dramatically%20from%20county%20to%20county [Accessed 1 Feb. 2025].

Kelly, J.T. et al. (2020). *Modifiable lifestyle factors for primary prevention of chronic kidney disease: a systematic review and meta-analysis.* Journal of the American Society of Nephrology, September. [online] Available at: https://doi.org/10.1681/ASN.2020030384 [Accessed 22 Feb. 2025].

Krishnakumar Thankappan (2024). *Sample Size in Machine Learning - Krishnakumar Thankappan - Medium.* [online] Medium. Available at: https://medium.com/@drkrishnakumart_51860/sample-size-in-machine-learning-74d5aa6eb321 [Accessed 21 Feb. 2025].

Krishnamurthy, S., KS, K., Dovgan, E., Luštrek, M., Gradišek Piletič, B., Srinivasan, K., Li, Y.-C. (Jack), Gradišek, A. and Syed-Abdul, S. (2021). *Machine Learning Prediction Models for Chronic Kidney Disease Using National Health Insurance Claim Data in Taiwan.* Healthcare, 9(5), p.546. [online] Available at: https://doi.org/10.3390/healthcare9050546 [Accessed 22 Feb. 2025].

Langote, M., Saratkar, S., Kumar, P., Verma, P., Puri, C., Gundewar, S. and Gourshettiwar, P. (2024). *Human–computer interaction in healthcare: Comprehensive review.* AIMS Bioengineering, 11(3), pp.343–390. [online] Available at: https://doi.org/10.3934/bioeng.2024018 [Accessed 22 Feb. 2025].

Lewis, S. (2024). *Canadians faced longest ever health-care wait times in 2024, study finds.* [online] Available at: https://nationalpost.com/news/canada/canadians-faced-long [Accessed 1 Feb. 2025].

Luong, K. (2024). *Challenges of AI Integration in Healthcare.* [online] Ominext.com. Available at: https://www.ominext.com/en/blog/challenges-of-ai-integration-in-healthcare [Accessed 14 Feb. 2025].

Naber, T. and Purohit, S. (2021). *Chronic Kidney Disease: Role of Diet for a Reduction in the Severity of the Disease.* Nutrients, 13(9), p.3277. [online] Available at: https://doi.org/10.3390/nu13093277 [Accessed 22 Feb. 2025].

Neale, E.P., Rosario, V.D., Probst, Y., Beck, E., Tran, T.B. and Lambert, K. (2023). *Lifestyle Interventions, Kidney Disease Progression, and Quality of Life: A Systematic Review and Meta-analysis.* Kidney Medicine, 5(6), pp.100643–100643. [online] Available at: https://doi.org/10.1016/j.xkme.2023.100643 [Accessed 22 Feb. 2025].

Office (2022). *Use of Online Tracking Technologies by HIPAA Covered Entities and Business Associates.* [online] HHS.gov. Available at: https://www.hhs.gov/hipaa/for-professionals/privacy/guidance/hipaa-online-tracking/index.html?utm_source=chatgpt.com [Accessed 22 Feb. 2025].

Pankaj Chittora, Sandeep Chaurasia, Chakrabarti, P., Gaurav Kumawat, Chakrabarti, T., Leonowicz, Z., Jasinski, M., Jasinski, L., Radomir Gono, Elzbieta Jasinska and Vadim Bolshev (2021). *Prediction of Chronic Kidney Disease - A Machine Learning Perspective.* IEEE Access, 9, pp.17312–17334. [online] Available at: https://doi.org/10.1109/access.2021.3053763 [Accessed 22 Feb. 2025].

Patwardhan, M.B., Kawamoto, K., Lobach, D., Patel, U.D. and Matchar, D.B. (2009). *Recommendations for a Clinical Decision Support for the Management of Individuals with Chronic Kidney Disease.* Clinical Journal of the American Society of Nephrology, 4(2), pp.273–283. [online] Available at: https://doi.org/10.2215/cjn.02590508 [Accessed 22 Feb. 2025].

Ross, C. (2018). *IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show.* [online] Available at: https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/ [Accessed 1 Feb. 2025].

Saif, D., Sarhan, A.M. and Elshennawy, N.M. (2024). *Early prediction of chronic kidney disease based on ensemble of deep learning models and optimizers.* Journal of Electrical Systems and Information Technology, 11(1). [online] Available at: https://doi.org/10.1186/s43067-024-00142-4 [Accessed 22 Feb. 2025].

Sandri, E., Larumbe, E.C., Capoferri, M., Olmedo, G.C., Werner, L.U. and M Jesús Vega-Bello (2024). *Socio-demographic determinants of dietary choices and their impact on health in Spanish adults.* Frontiers in Public Health, 12. [online] Available at: https://doi.org/10.3389/fpubh.2024.1417925 [Accessed 22 Feb. 2025].

Saye, T. (2024). *Data-Driven Personal Training.* [online] Ptdistinction.com. Available at: https://www.ptdistinction.com/blog/data-driven-personal-training?utm_source=chatgpt.com [Accessed 22 Feb. 2025].

ul Abideen, Z., Khan, T.A., Ali, R.H., Ali, N., Baig, M.M. and Ali, M.S. (2022). *DocOnTap: AI-based disease diagnostic system and recommendation system.* [online] IEEE Xplore. Available at: https://doi.org/10.1109/ICET56601.2022.10004621 [Accessed 22 Feb. 2025].

Vairale, V.S. and Shukla, S. (2018). *Recommendation framework for diet and exercise based on clinical data: A systematic review.* Lecture Notes on Data Engineering and Communications Technologies, pp. 333–346. [online] Available at: https://doi.org/10.1007/978-981-10-7641-1_29 [Accessed 22 Feb. 2025].

Villanego, F., Naranjo, J., Vigara, L.A., Cazorla, J.M., Montero, M.E., García, T., Torrado, J. and Mazuecos, A. (2020). *Impacto del ejercicio físico en pacientes con enfermedad renal crónica: revisión sistemática y metaanálisis.* Nefrología, 40(3), pp.237–252. [online] Available at: https://doi.org/10.1016/j.nefro.2020.01.002 [Accessed 22 Feb. 2025].

World Health Organization (2024). *The top 10 causes of death.* [online] Available at: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death [Accessed 31 Jan. 2025].

World Health Organization (2024). *Noncommunicable diseases.* [online] Available at: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [Accessed 1 Feb. 2025].