RE-2022-137155-plag-report

by Research Experts - Turnitin Report

Submission date: 07-Jun-2023 02:36PM (UTC-0400)
Submission ID: 2111209869
File name: RE-2022-137155.pdf (1,010.34K)
Word count: 8775
Character count: 51021

CHRONIC KIDNEY DISEASE DETECTION USING MACHINE LEARNING

A Project Report Submitted
in Partial Fulfillment of the Requirements
for the Degree of

Bachelor of Technology

in

Computer Science and Business System

by

Mr. Vikash Kumar Ramesh Goyat PRN No.:1943110091
Miss Sudiksha Ravale PRN No.:2043110285
Miss Ishika Verma PRN No.:1943110015

Under the guidance of

Prof. Sumita Kumar

COMPUTER SCIENCE AND BUSINESS SYSTEM
BHARATI VIDYAPEETH (D.U.)
DEPARTMENT OF ENGINEERING AND TECHNOLOGY, OFF CAMPUS, NAVI MUMBAI

Academic Session 2022-23

UNDERTAKING

We declare that the work presented in this project report titled "CHRONIC KIDNEY DISEASE DETECTION USING MACHINE LEARNING", submitted to the Computer Science and Business System Department, Bharati Vidyapeeth Deemed to be University, Pune, Department of Engineering and Technology, Off Campus, Navi Mumbai, for the award of the ***Bachelor of Technology*** degree in ***Computer Science and Business System***, is our original work. We have not plagiarized or submitted the same work for the award of any other degree. In case this undertaking is found incorrect, We accept that my degree may be unconditionally withdrawn.

May 24, 2023
Navi Mumbai

Mr. Vikash Kumar
Ramesh Goyat PRN
No.:1943110091
Miss Sudiksha Ravale
PRN No.:2043110285
Miss Ishika Verma PRN
No.:1943110015

Bharati Vidyapeeth Deemed to be University
Department of Engineering and Technology, Offcampus, Navi Mumbai

CERTIFICATE

Certified that the work contained in the project report titled "***Predicting Chronic kidney Disease using Machine Learning***", by the following students:

Mr. Vikash Kumar Ramesh Goyat PRNNo.1943110091

Miss Sudiksha Ravale PRN No.:2043110285

Miss Ishika Verma PRN No.:1943110015

has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

____

(Project Guide)
Prof. Sumita Kumar
Associate Professor
Computer Science and Business System

____

Dr. Shabir Ali
Head of Department,
Computer Science and Business System
DET, Navi Mumbai

____

Dr. Mohan Awasthy
Principal
DET, Navi Mumbai

Date of Certificate: ______________________________

Bharati Vidyapeeth Deemed
to be University
Department of Engineering and Technology, Offcampus,
Navi Mumbai

APPROVAL CERTIFICATE

This project title "***Predicting Chronic kidney Disease using Machine Learning*** " by the following students:

Mr. Vikash Kumar Ramesh Goyat PRNNo.1943110091

Miss Sudiksha Ravale PRN No.:2043110285

Miss Ishika Verma PRN No.:1943110015

has been approved for the degree of *Bachelor of Technology* in *Computer Science and Business System* from Department of Engineering and Technology, Off campus, Navi Mumbai, Bharati Vidyapeeth (Deemed to be University), Pune.

Examiners:

External Examiner Name and Sign

Internal Examiner Name and Sign

Date of Approval: ____________________

Acknowledgements

We would like to express my sincere gratitude to HoD **Dr. Shabir Ali** of **Department of Computer Science and Business Systems** of Institute **”Bharati Vidyapeeth (D.U.), Department of Engineering and Technology, Off campus, Navi Mumbai”** for their valuable support and guidance during this project. His dedication towards providing quality education, state-of-the-art infrastructure, and research opportunities has been instrumental in enabling us to undertake this project. The Institute’s resources, facilities, and faculty members have provided us with an excellent platform to learn, grow and explore our potential.

We would like to express our sincere gratitude to our Supervisor **Prof. Sumita Kumar** for his/her invaluable contributions to this project.

Prof. Sumita Kumar guidance, support, and mentorship have been critical in helping us to develop a deep understanding of the subject matter and to undertake this project with confidence. His/her constructive feedback, attention to detail, and commitment to excellence have inspired us to strive for the highest standards of quality and professionalism.

Throughout the project, Prof. Sumita Kumar provided us with her extensive knowledge, expertise, and insights, which helped us to navigate through the challenges and make informed decisions. His/her collaborative and inclusive approach towards learning has fostered a culture of innovation and creativity, which has been vital in shaping our ideas and perspectives.

Moreover, We would like to thank Prof. Sumita Kumar for her generosity in sharing his/her time, resources, and expertise with us. His/her unwavering support and encouragement have been instrumental in helping us to complete this project successfully.

Finally, We would like to express my deep appreciation to Prof. Sumita Kumar for her unwavering support and mentorship throughout the project. We are truly fortunate to have had the opportunity to work with him/her, and we will always cherish this experience.

Once again, thank you, Prof. Sumita Kumar, for your invaluable contributions to this project. We deeply appreciate all that you have done for us, and we are grateful for your guidance, support, and mentorship.

This Dissertation is Dedicated

To Mrs Sumita Kuma , whose support, guidance, and inspiration have been instrumental in making this research possible. Mrs Sumita Kumar has been a constant source of encouragement and motivation throughout the journey of this dissertation. Their unwavering commitment to our project has been critical in shaping our ideas and perspectives, and their leadership and mentorship have been invaluable in navigating the challenges and complexities of the research process.

Abstract

Chronic kidney disease (CKD), also known as chronic renal disease. Chronic kidney disease involves conditions that damage your kidneys and decrease their ability to keep you healthy. You may develop complications like high blood pressure, anemia (low blood count), weak bones, poor nutritional health and nerve damage. Early detection and treatment can often keep chronic kidney disease from getting worse. Data Mining is the term used for knowledge discovery from large databases. The task of data mining is to make use of historical data, to discover regular patterns and improve future decisions, follows from the convergence of several recent trends: the lessening cost of large data storage devices and the ever-increasing ease of collecting data over networks; the expansion of robust and efficient machine learning algorithms to process this data; and the lessening cost of computational power, enabling use of computationally intensive methods for data analysis. Machine learning, has already created practical applications in such areas as analyzing medical science outcomes, detecting fraud, detecting fake users etc. Various data mining classification approaches and machine learning algorithms are applied for prediction of chronic diseases. The objective of this research work is to introduce a new decision support system to predict chronic kidney disease. The aim of this work is to detect CKD using Support vector machine (SVM) classifier on the basis of its accuracy, precision and execution time.

Keywords – SVM , CNN , Confusion Matrix , Machine Learning , Data Mining

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

We all knows, that Kidney is essential organ in human body. Which has main functionalities like excretion and osmoregulation. In simple words we can say that all the toxic and unnecessary material from the body is collected and thrown out by kidney and excretion system. It is a dangerous disease of the kidney which produces gradual loss in kidney functionality. CKD is a slow and periodical loss of kidney function over a period of several years. If CKD is not detected and cured in early stage then patient can show following Symptoms: Blood Pressure, anaemia, weak bones, poor nutrition health and nerve damage, decreased immune response because at advanced stages dangerous levels of fluids, electrolytes, and wastes can build up in your blood and body. Hence it is essential to detect CKD at its early stage but it is unpredictable as its Some people have no symptoms at all so machine learning can be helpful in this problem to predict that the patient has CKD or not. Machine learning does it by using old CKD patient data to train predicting model. and determine the chronic kidney disease.

1.1 Problem Statement

The sudden failure of kidneys to perform the function. It is a serious condition and mainly occurs during the last stage of the disease or in time of treatment. This study assists doctors in exploring preventive measures for CKD through early diagnosis using machine learning techniques.

1.2 Scope

The scope of CKD is aims to provide a platform for further research findings for both clinical science and basic research and covers the entire spectrum of nephrological disorders. Clinical trials, observational studies, and other original clinical research work related to kidney diseases are the focus of the Clinical Science Section. Moreover, important studies on hemodialysis and peritoneal dialysis are also considered. Invited review articles and editorials by renowned experts and high-quality rapid communications round off the contents of the journal.

1.3 Motivation

Chronic kidney disease is a disease which grows slowly in the body and it affects deeply. This disease can harm people of any age. If this disease is not detected at an early stage, then it can cause a lot of trouble and the patient might go at the higher stage 12 where the condition of the patient can get critical. Therefore, this motivates us to work on a system which can help detect this disease in patients at an early stage .

1.4 Objectives

The proposed system deals with the prediction of chronic disease from the clinical data. The healthcare generates large data, so it is necessary to collect this data and effectively use it for analysis, prediction, and treatment. A classification model draws some conclusion from observed values. In classification model one or more inputs are used to predict the value of one or more outcomes

Chapter 2

Literature Review

1. Prediction of Chronic Kidney Disease using Adaptive Hybridized Deep Convolutional Neural Network on the Internet of Medical Things Platform.

Author: Guozhen Chen, Chenguang Ding

Abstract: Chronic Kidney disease is a severe lifelong condition caused either by renal disease or by impaired functions of the kidneys. In the present area of research, Kidney cancer is one of the deadliest and crucial importance for the survival of the patients' diagnosis and classification. Early diagnosis and proper therapy can stop or delay the development of this chronic disease into the final stage where dialysis or renal trans plantation is the only way of saving the life of the patient. The development of automated tools to accurately identify subtypes of kidney cancer is, therefore, an urgent challenge in the recent past. In this paper, to examine the ability of various deep learning methods an Adaptive hybridized Deep Convolutional Neural Network (AHDCNN) has been proposed for the early detection of Kidney disease efficiently and effectively. Classification technology efficiency de pends on the role of the data set. To enhance the accuracy of the system by reducing the feature dimension classification an algorithm model has been developed using CNN. These high-level properties help to build a supervised tissue classifier that discriminates between the two types of tissue.

2. Diagnostic Decision Support System of Chronic Kidney Disease Using Support Vector Machine.

Author Mubarik Ahmad, Vitri Tund jungsari

Abstract: Kidney disease or commonly known as kidney failure is a condition when the renal function is declining that could result in the inability of the kidneys to perform their duties. Kidney disease patients have the potential to get into the chronic phase. Chronic kidney disease is a decrease in kidney function gradually during the three months which resulted in the cessation of kidney function in total. The purpose of this development is a decision support system for a doctor in diagnosing of the kidney disease patients. The system displays the results of predicting whether patients with renal disease have entered a phase of chronic kidney disease or not. The methodology of this study con sists of two main phases: classification modelling and system development. Classification modelling consists of data collection, data preparation, data grouping, classification, rules extraction. System development was based on the extracted rules before. This study resulted in a system that can detect a chronic condition of kidney disease based on several factors with an accuracy of 98.34.

3. Early Prediction of Chronic Kidney Disease Using Ma chine Learning Supported by Predictive Analytics

Author: Ahmed J. Aljaaf, Dhiya Al-Jumeily

Abstract: Chronic Kidney Disease is a serious lifelong condition that induced by either kidney pathology or reduced kidney functions. Early prediction and proper treatments can possibly stop, or slow the progression of this chronic disease to end-stage, where dialysis or kidney trans plantation is the only way to save patient's life. In this study, we examine the ability of several machine-learning methods for early pre diction of Chronic Kidney Disease. This mat ter has been studied widely; however, we are supporting our methodology by the use of predictive analytics, in which we examine the relationship in between data parameters as well as with the target class attribute. Predictive analytics enables us to introduce the optimal subset of parameters to feed machine learn ing to build a set of predictive models. This study starts with 24 parameters in addition to the class attribute, and ends up by 30% of them as ideal sub set to predict Chronic Kidney Dis ease. A total of 4 machine learning based classifiers have been evaluated within a super vised learning setting, achieving highest performance outcomes of AUC 0.995, sensitivity 0.9897, and specificity 1. The experimental procedure concludes that advances in ma chine learning, with assist of predictive analytics, represent a promising setting by which to recognize intelligent solutions, which in turn prove the ability of predication in the kidney disease domain and beyond.

4. Analysis of Chronic Kidney Disease Dataset by Applying Ma chine Learning Methods

Author: Yedilkhan Amir galiyev, Shahriar Shamilu ulu.

Abstract: Currently, there are many people in the world suffering from chronic kidney diseases worldwide. Due to the several risk factors like food, environment and living standards many people get diseases suddenly without understanding of their condition. Diagnosing of chronic kidney diseases is generally invasive, costly, time-consuming and often risky. That is why many patients reach late stages of it without treatment, especially in those countries where the resources are limited. Therefore, the early detection strategy of the disease remains important, particularly in developing countries, where the diseases are generally diagnosed in late stages. Finding a solution for above mentioned problems and riding out from disadvantages became a strong motive to con duct this study. In this research study, the effects of using clinical features to classify patients with chronic kidney disease by using support vector machines algorithm is investigated. The chronic kidney disease dataset is based on clinical history, physical examinations, and laboratory tests. Experimental results showed over 93% of success rate in classifying the patients with kidney diseases based on three performance metrics i.e., ac curacy, sensitivity and specificity.

5. Extraction of Action Rules for Chronic Kidney Disease using Naïve Bayes Classifier.

Author: Dr. Uma N Dul hare, Mo hammad Ayesha

Abstract: Chronic kidney disease (CKD), also known as chronic renal disease, which is progressive loss in kidney function over a period of months or years. It is defined by the presence of kidney damage or decreased glomerular filtration rate (GFR). The estimated prevalence of CKD is about 9-13adult population. Individuals with CKD have a far greater likelihood of cardiovascular death than progression to end-stage renal disease. CKD is more
prevalent in patients with CVD or with CVD related risk factors, such as hypertension, diabetes mellitus, dyslipidaemia, and metabolic syndrome In proposed work, we are not only extracting action rules based on stages but also predicting CKD by using naive bayes with One attribute selector which helps to prevent the advancing of chronic renal dis ease to further stages.

6. Cellular-Level Structure Imaging with Micro-optical Coherence Tomography (µOCT) for
Kidney Disease Diagnosis.

Author: Chi Hu, Xiaojun Yu, Qianshan Ding

Abstract: Chronic kidney disease (CKD) is one of the public health threats around the world, which may cause serious health problems like cardiovascular disease, kidney failure, or even more serious as premature death. Although CKD usually could be managed by general internists, such a way of treatment can be applied only when significant symptoms appear, which is very slow. It has also been reported that CKD could be characterized by means of glomeruli, and classified by the stages of disease severity for early treat ment. However, due to lack of reliable method to detect the cellular-level microstructures for dis ease severity characterization, the diagnosis is troublesome, and thus, the treatments might be delayed while the best treatment time

could be missed. For early detection of CKD, it is imperative to develop reliable tools to detect and characterize the disease at an early stage with minimal or non-invasiveness. For this research, the micro-optical coherence tomography (OCT) was assessment it feasibility as a cellular level structure imaging tool in kidney disease diagnosis at an early stage. Specifically, by measuring the number of glomeruli within a volumetric kidney tissue, a new diagnostic criterion is also established. Imaging results of the kidney specimens as compared their corresponding histology show that the cellular level glomeruli structures could be identified clearly, and as a basic functional unit of kidney, it could be utilized as a reliable parameter to access the severity of the CKD.

7. Texture Analysis of Ultrasound Images of chronic kidney disease.

Author: Fadil Iqbal1, Aruna S. Pallewatte.

Abstract: Chronic Kidney Disease of unknown aetiology (CKDu) is a prevalent disease in the North Central Province of Sri Lanka. Towards the latter stages of the disease, kidney function fails by 80 interstitial fibrosis is formed and grows as the disease progresses. The cause of the disease remains elusive and early detection is vital to arrest the progressive decline of kidney function. The objective of this study is to construct a computer program to perform texture analysis on ultrasound kidney images and extract various features that can be used to distinguish between normal and diseased kidney patients. The computer program was developed using MATLAB and a user interface was created to perform mathematical operations such as: Fourier analysis to extract Root Mean Square and First Moment values and Grey Level Co-occurrence Matrix (GLCM) to extract Homogeneity and Sum Average values. A sample of ultrasound images were taken from 32 patients. Region of interest (ROI) selection was performed on entire kidney, cortex region and white (renal medulla or renal sinus) region separately. Among these methods Root Mean Square values over the entire kidney (p=0.03) and cortex region (p=0.0049) gave significant results in distinguishing between normal and diseased kidneys.

8. Characterizing volumes of kidney segments in Streptozotocin induced diabetic rat model
utilizing 4D contrast enhanced ultra sound.

Author: Kennita A. Johnson, A. Gloria Nyankima, Paul A.

Abstract: Diabetic Kidney Disease is a disease that if left uncontrolled may eventually lead to end stage kidney disease. Clinicians utilize biomarkers, such as albuminuria and serum creatinine, to identify diabetic populations at risk for kidney disease, but these markers often lag behind histologic disease. Contrast enhanced ultrasound (CEUS) presents a potential tool to identify at risk patients, begin early intervention and prevent development and progression of kidney disease. Utilizing 4D CEUS, we observed the progressive effect of diabetes in a rodent streptozotocin (STZ) model, over a 12-week

period. The following describes the experimental proto col and image process for characterization of the kidneys. Image datasets estimated total volume changes accurately in comparison to water displacement volume measurements. Treated animals displayed an increase in total volume, as expected. Individual cortical and medullary volumes were estimated using the same datasets and showed a greater increase in medullary than cortical volumes. Further investigation is required to validate CEUS as a tool for early DKD diagnosis.

9. Chronic Kidney Disease Prediction and Recommendation of Suitable Diet plan by using
Machine Learning.

Author: Akash Maurya, Rahul Wable, Rasika Shinde

Abstract: Chronic kidney disease (CKD) is a type of kidney disease in which there is gradual loss of kidney function over a period of months or years. Prediction of this dis ease is one of the most important problems in medical fields. So automated tool which will use machine learning techniques to determine the patient's kidney condition that will be helpful to the doctors in prediction of chronic kidney disease and hence better treatment. The proposed system extracts the features which are responsible for CKD, then machine learning process can automate the classification of the chronic kidney disease in different stages according to its severity. The objective is to use machine learning algorithm and suggest suitable diet plan for CKD patient using classification algorithm on medical test records. Diet recommendation for patient will be given according to potassium zone which is calculated using blood potassium level to slow down the progression of CKD.

Chapter 3

Requirement Analysis

3.1 Software Requirement

The prerequisites software and libraries for the sign language project are:

- **Programming Language**: You will need a programming language to implement your CKD prediction model. Popular options for machine learning projects include Python and R. Python, in particular, has a rich ecosystem of libraries and frameworks that make it well-suited for machine learning tasks.
- **Integrated Development Environment (IDE)**: An IDE provides a comprehensive development environment for coding, debugging, and running your project. Some popular IDEs for Python include PyCharm, Spyder, and Jupyter Notebook. For R, you can use RStudio or Jupyter Notebook as well.
- **Machine Learning Libraries**: You will need specific machine learning libraries to implement the predictive models for CKD. Here are some widely used libraries in Python:
- **scikit-learn**: This is a popular machine learning library that provides various algorithms and tools for classification, regression, and other tasks. It includes implementations of SVM, Random Forest, Decision Tree, KNN, and many other algorithms.
- **pandas**: This library provides data structures and data analysis tools for manipulating and preprocessing datasets. It is commonly used for cleaning and preparing the CKD dataset before feeding it into the machine learning models.
- **NumPy**: NumPy is a fundamental library for numerical computations in Python. It provides support for multidimensional arrays and mathematical functions, which are used extensively in machine learning tasks.
- **Jupyter :** it is widely used and powerful open-source integrated development environment (IDE) that is particularly useful in the fields of data science and machine learning. It provides a platform for users to create and share documents that contain live code, equations, visualizations, and narrative text. This makes it an excellent tool for interactive exploration and prototyping in a wide variety of applications.
- **matplotlib and seaborn**: These libraries are used for data visualization, enabling

you to create informative plots and charts to understand the data and evaluate the performance of your models.

- These libraries can be installed using package managers like pip or conda.
- **Database Management System (DBMS)**: If you are working with a large dataset or planning to store data in a database, you might need a DBMS like MySQL, PostgreSQL, or SQLite to manage and query the data efficiently.
- **Version Control**: It is highly recommended to use version control software, such as Git, to track changes in your project and collaborate with team members effectively.
- **Documentation and Reporting Tools**: Tools like Jupyter Notebook, Markdown, or LaTeX can help you document your project, explain your methodologies, and present your findings in a clear and organized manner.

3.2 Hardware Requirements

Here are some general hardware requirements

- **Computer System**: You will need a computer system or server with sufficient processing power and memory capacity to handle the data processing and modelling tasks involved in CKD prediction. The exact hardware specifications will depend on the size of your dataset and the complexity of the machine learning algorithms you are using. A modern multi-core processor (e.g., Intel Core i5 or higher) and a minimum of 8 GB RAM are recommended.

- **Storage**: Sufficient storage space is necessary to store the dataset and any intermediate or final model files generated during the project. The storage capacity required will depend on the size of your dataset and any additional data that needs to be stored. Ensure you have enough storage available, preferably in the form of a solid-state drive (SSD) for faster read/write operations.

- **Graphics Processing Unit (GPU)**: While not strictly necessary, using a GPU can significantly speed up the training and inference processes, especially for complex machine learning models. GPUs excel at parallel processing, which is highly beneficial for computationally intensive tasks like training deep learning models. If you plan to work with large datasets or employ advanced deep learning models, consider investing in a dedicated GPU.

- **Internet Connectivity**: Internet connectivity is crucial for downloading necessary software libraries, updates, and accessing relevant research papers or resources. Additionally, if your project involves collecting data from external sources or utilizing cloud-based services for model training or deployment, a stable internet connection is necessary.

- **Backup and Data Redundancy**: It is essential to have a reliable backup system in place to prevent data loss or corruption. Regularly back up your project files, including the dataset, code, and any other project-related data, to an external storage device or a cloud-based backup service. This ensures that you can recover your work in case of hardware failures or other unforeseen circumstances.

- **Peripherals**: Standard computer peripherals like a keyboard, mouse, and display monitor are necessary for interacting with the system and visualizing project results. Additionally, if you plan to work with physical data collection devices (e.g., sensors), ensure you have the required hardware interfaces and adapters to connect them to your system.

Chapter 4

Report on Present Investigation

4.1 Proposed System

Proposed system aims to automate the process of diagnosing chronic kidney disease by leveraging data analysis techniques and machine learning algorithms. It streamlines the workflow from data import to prediction and performance analysis, providing a comprehensive approach to assist medical professionals in the diagnosis and management of the disease.

system checks if the required libraries for data analysis and machine learning are imported. If not, it prompts to install them. Next, the system checks if the dataset for chronic kidney disease is imported. If not, it suggests loading the dataset from a reliable source.Once the dataset is available, the system performs data preprocessing to ensure its quality. This step involves cleaning the data, handling missing or null values, and performing any necessary transformations or normalization. If there are missing values in the dataset, the system offers the option of using K-Nearest Neighbors (KNN) imputation to fill in those missing values. Next, the system provides the option to apply classification algorithms to predict the presence or absence of chronic kidney disease based on the available data. If classification is not needed, this step can be skipped. The system then suggests splitting the dataset into a training set and a test set for model evaluation. This allows for assessing the performance of the classification algorithms on unseen data. If chosen, the system applies several algorithms for prediction, including Support Vector Machine, Random Forest, Decision Tree, and K-Nearest Neighbor. Each algorithm may have its own strengths and weaknesses, and applying multiple algorithms helps to evaluate their comparative performance. After making predictions using the algorithms, the system offers the option to analyze the performance of the predictions. This analysis helps in assessing the accuracy, precision, recall, and other relevant metrics to gauge the effectiveness of the chosen algorithms

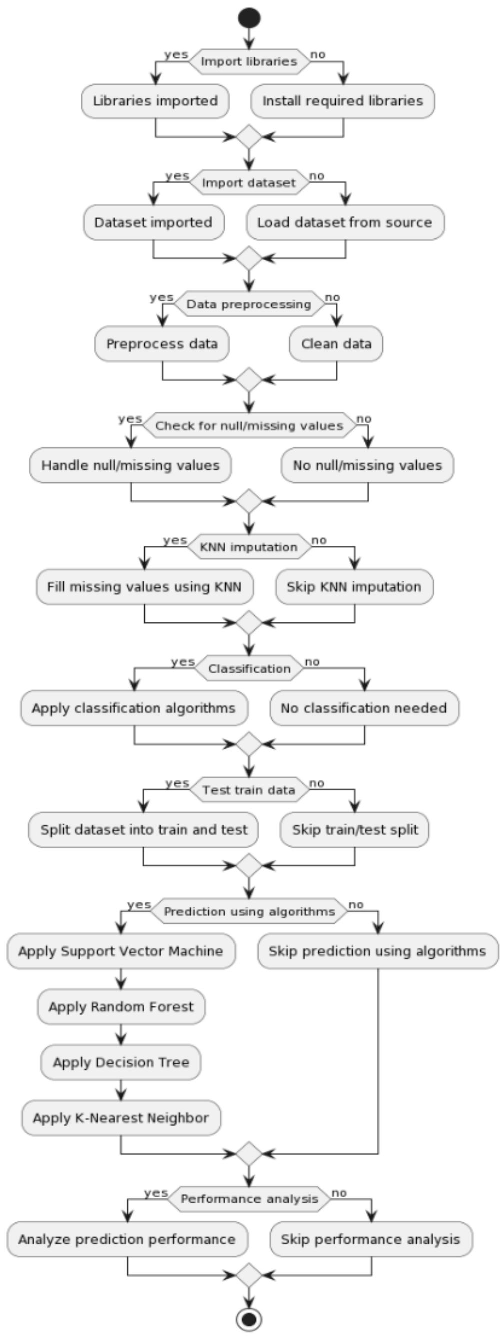

Figure 4.1: Proposed System Flowchart

4.1.1 Block Diagram of Proposed System

Input system checks if the required libraries for data analysis and machine learning are imported. If not, it prompts to install them. Next, the system checks if the dataset for chronic kidney disease is imported. If not, it suggests loading the dataset from a reliable source.

Once the dataset is available, the system performs data preprocessing to ensure its quality. This step involves cleaning the data, handling missing or null values, and performing any necessary transformations or normalization.

If there are missing values in the dataset, the system offers the option of using K-Nearest Neighbors (KNN) imputation to fill in those missing values.

Next, the system provides the option to apply classification algorithms to predict the presence or absence of chronic kidney disease based on the available data. If classification is not needed, this step can be skipped.

The system then suggests splitting the dataset into a training set and a test set for model evaluation. This allows for assessing the performance of the classification algorithms on unseen data.

If chosen, the system applies several algorithms for prediction, including Support Vector Machine, Random Forest, Decision Tree, and K-Nearest Neighbor. Each algorithm may have its own strengths and weaknesses, and applying multiple algorithms helps to evaluate their comparative performance.

After making predictions using the algorithms, the system offers the option to analyze the performance of the predictions. This analysis helps in assessing the accuracy, precision, recall, and other relevant metrics to gauge the effectiveness of the chosen algorithms.

4.1 Block Diagram of Proposed System

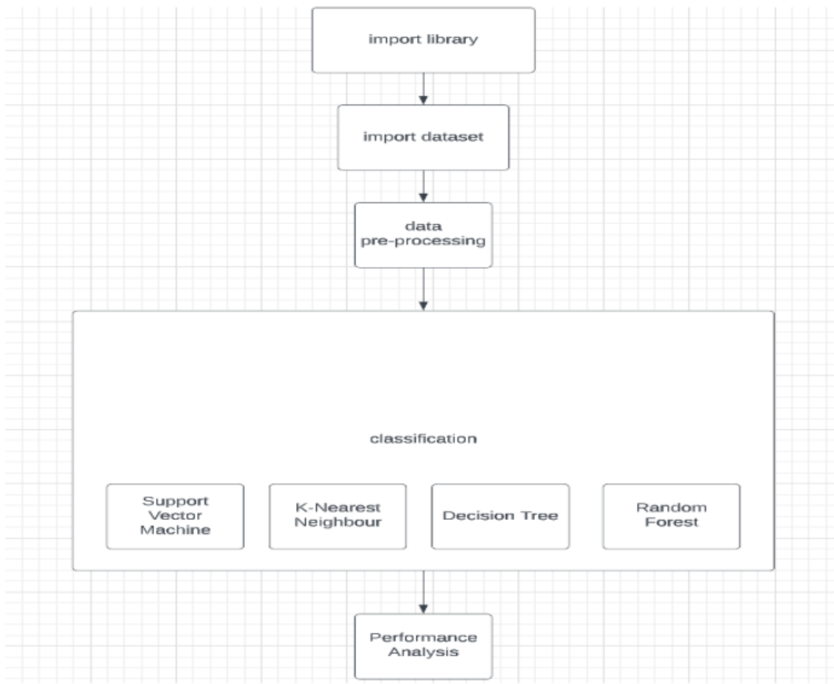

Figure 4.2: Block Diagram of Proposed System

- The proposed system for predicting chronic kidney disease (CKD) incorporates four machine learning algorithms: Support Vector Machine (SVM), Random Forest, Decision Tree, and K-Nearest Neighbours (KNN). The block diagram of the system illustrates the flow of data and operations involved in the prediction process.

- **Data Collection**: The system begins with the collection of relevant data, which typically includes clinical and demographic variables associated with CKD. These variables may include age, gender, blood pressure, blood glucose levels, urinary markers, and other medical history indicators. The dataset serves as the input for the predictive models.

- **Pre-processing**: The collected data undergoes pre-processing to ensure its quality and suitability for the machine learning algorithms. This step involves data cleaning, handling missing values, normalization, and feature scaling. Pre-processing helps to eliminate noise, standardize the data, and improve the performance and accuracy of the models.

- **Feature Selection/Engineering**: In this step, feature selection techniques may be employed to identify the most relevant variables that contribute significantly to

CKD prediction. Alternatively, feature engineering techniques can be used to create new features based on domain knowledge or transformation of existing variables. This process helps to reduce dimensionality and focus on the most informative features.

- **Model Training**: The pre-processed and selected features are used to train the four machine learning models: SVM, Random Forest, Decision Tree, and KNN. Each algorithm learns from the input data to create a predictive model that can classify individuals as either having CKD or not. The training process involves adjusting the model parameters to optimize its performance based on the training data.

- **Model Evaluation**: The trained models are then evaluated using validation data to assess their performance and generalization capability. Standard evaluation metrics such as accuracy, precision, recall, and F1-score are calculated to measure the models' predictive accuracy and effectiveness in identifying CKD cases. This evaluation helps in selecting the best-performing model for further analysis.

- **Model Selection**: Based on the evaluation results, the best-performing model, which in this case is Random Forest, is selected for CKD prediction. Random Forest demonstrates superior performance compared to the other three models (SVM, Decision Tree, and KNN) in terms of accuracy and other evaluation metrics.

- **Prediction**: The selected Random Forest model is then used to predict CKD in new, unseen data. The system takes the input of clinical and demographic variables for an individual and passes it through the trained Random Forest model. The model evaluates the input features and generates a prediction of whether the individual is likely to have CKD or not.

- **Result Output**: The final output of the system is the predicted CKD status for the individual based on the Random Forest model. This prediction can help healthcare professionals make informed decisions regarding the individual's risk of CKD, initiate appropriate interventions, and provide personalized care.

4.1.2 UML Diagrams

1. **Use Case Diagram** Use case between explanation and analysis of requirements to represent system performance. Use case description of the function of the system which gives visible results for the actor. Identifying actors and their problem cases by lecturing on the boundaries of the system, rep-resenting the work done by them and the whole environment. Actors are outside the system, while cases are within the system. The use case describes the system as seen from the example of the actor's behavior. It describes the work provided by the system as a set of events that provide visible results for the actor.

   In this updated diagram, we added use cases for importing libraries and importing datasets. After importing the dataset, the data preprocessing step is performed, followed by the classification step. The classification step includes machine learning algorithms like Random Forest, Decision Tree, SVM, and KNN. Finally, the performance analysis is conducted.

- Import Libraries: This use case represents the step where the necessary libraries or packages are imported into the system. These libraries typically include machine learning frameworks, data manipulation libraries, and other dependencies required for the prediction process.

- Import Dataset: This use case involves importing the dataset containing patient data for chronic kidney disease prediction. The dataset may include features such as age, blood pressure, serum creatinine levels, etc. This data will be used for training and testing the machine learning models.

- Data Preprocessing: This use case focuses on the preprocessing of the imported dataset. It includes tasks such as handling missing values, scaling features, encoding categorical variables, and any other necessary data transformations required to prepare the data for the subsequent steps.

- Classification: This use case represents the overarching step of classification, where different machine learning algorithms are used to build predictive models for chronic kidney disease. This step involves training the models on the preprocessed data and using them to make predictions.

- Random Forest, Decision Tree, SVM, KNN: These use cases represent the specific machine learning algorithms utilized for chronic kidney disease prediction. Each algorithm, such as Random Forest, Decision Tree, Support Vector Machines (SVM), and k-Nearest Neighbors (KNN), may have different characteristics and strengths in capturing patterns in the data.

- Performance Analysis: This use case involves evaluating and analyzing the performance of the trained models. Various performance metrics, such as accuracy, precision, recall, and F1 score, can be calculated to assess how well the models are predicting chronic kidney disease. This analysis helps in understanding the effectiveness of different algorithms and selecting the best-performing one.

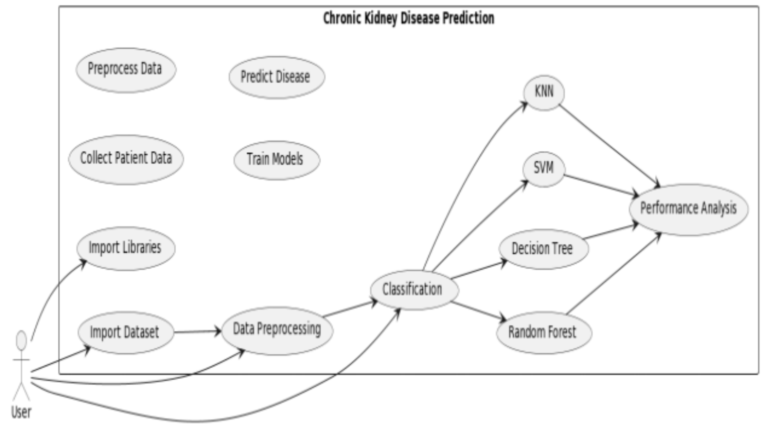

Figure 4.3: Use Case Diagram

2. **Sequence Diagram** The sequence diagram displays the time sequence of the objects participating in the interaction.
   In this diagram, the User interacts with the system by providing the Chronic Kidney Dataset. The Data Preprocessing component preprocesses the data, and the Machine Learning Algorithms component trains the models using SVM, KNN, Decision Tree, and Random Forest algorithms one by one. Each algorithm trains the model, predicts chronic kidney based on the provided patient data, and returns the prediction to the Machine Learning Algorithms component. Finally, the Machine Learning Algorithms component displays the predictions to the User.

- The User provides the Chronic Kidney Dataset to the Data Preprocessing component.
- The Data Preprocessing component preprocesses the provided data.
- The Machine Learning Algorithms component receives the preprocessed data and proceeds to train a model using the Support Vector Machine (SVM) algorithm.
- The Machine Learning Algorithms component sends the trained SVM model to the Chronic Kidney Prediction Model component.
- The Chronic Kidney Prediction Model component receives the trained SVM model and prepares to predict chronic kidney using the provided patient data.

- The Machine Learning Algorithms component requests the Chronic Kidney Prediction Model component to predict chronic kidney using the SVM model.
- The Chronic Kidney Prediction Model component performs the prediction using the SVM model and returns the prediction to the Machine Learning Algorithms component.
- Steps 3-7 are repeated for the K-Nearest Neighbors (KNN), Decision Tree, and Random Forest algorithms. Each algorithm is trained, used for prediction, and the prediction result is returned to the Machine Learning Algorithms component.
- The Machine Learning Algorithms component receives the predictions from all the algorithms.
- The Machine Learning Algorithms component displays the predictions to the User.

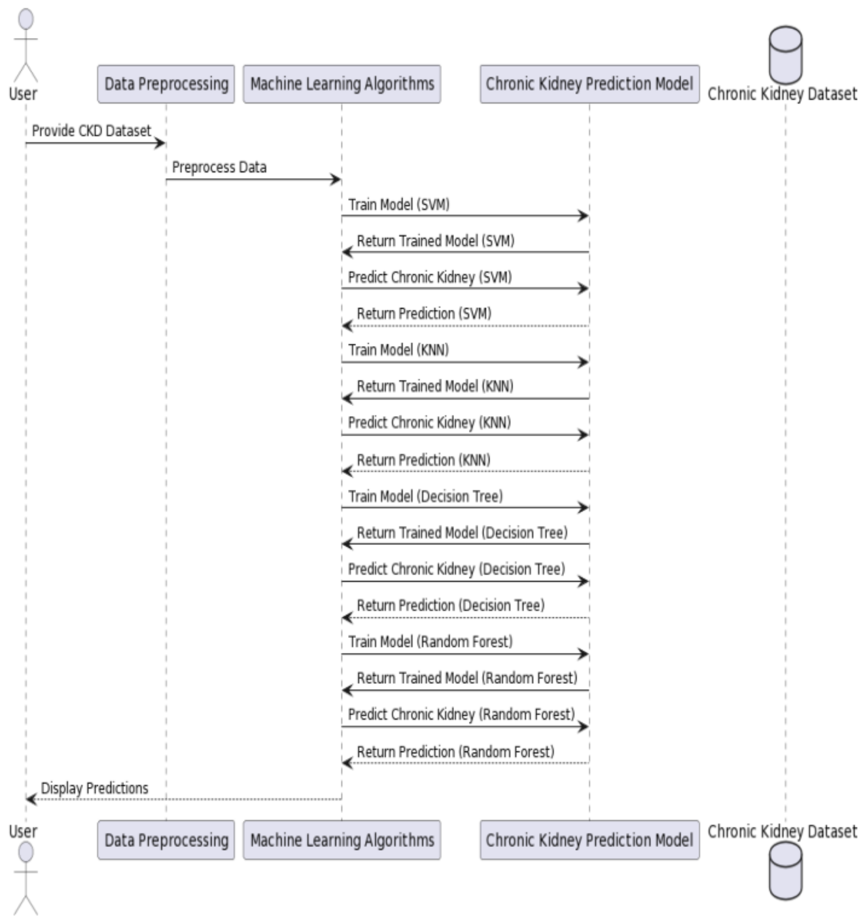

Figure 4.4: Sequence Diagram

3. **Activity Diagram** Object calling methods use messages and add new activation boxes on the second vertex to indicate the level of the next process.

The activity diagram begins with importing libraries and importing the dataset. Then, it proceeds to perform data preprocessing. Next, it checks for null values and replaces them with "?" as mentioned. Following that, it applies K-Nearest Neighbors (KNN) imputation to fill in the missing values. The classification phase involves testing and training the data, followed by prediction using different machine learning algorithms such as Support Vector Machine (SVM), Random Forest, Decision Tree, and K-Nearest Neighbor (KNN). Finally, it concludes with the performance analysis. As before, you can copy and paste this code into a PlantUML editor or use a PlantUML plugin to generate the corresponding diagram.

- Import Libraries: This step involves importing the necessary libraries or packages required for the prediction process. It ensures that all the required dependencies are available.

- Import Dataset: In this step, the dataset containing chronic kidney disease-related information is imported. This dataset serves as the foundation for the prediction model.

- Data Preprocessing: The dataset undergoes data preprocessing to prepare it for further analysis. This step involves tasks such as cleaning the data, handling missing values, transforming variables, and normalizing data if necessary.

- Check for Null Values: The dataset is checked for any null values or missing values. This is an important step to ensure the quality of the dataset and to determine if any data is incomplete.

- Replace Null Values: If null values are found in the dataset, they are replaced with a specific placeholder symbol, "?" in this case. This allows for consistent handling of missing values during the subsequent stages.

- K-Nearest Neighbors (KNN) Imputation: KNN imputation is used to fill in the missing values in the dataset. This algorithm calculates the values for missing data points based on the values of neighboring data points.

- Classification: The dataset is divided into training and testing sets. The training set is used to train the machine learning models, while the testing set is used to evaluate the performance of the models.

- Test Train Data: This step involves splitting the dataset into training and testing subsets. The training set is used to train the models, and the testing set is used to assess their accuracy.

- Prediction using Different Machine Learning Algorithms: Various machine learning algorithms, including Support Vector Machine (SVM), Random Forest, Decision

Tree, and K-Nearest Neighbor (KNN), are employed to predict chronic kidney disease based on the trained models.

- Performance Analysis: The performance of the prediction models is evaluated using different metrics such as accuracy, precision, recall, or F1 score. This step helps in understanding the effectiveness and reliability of the models. The activity diagram provides a visual representation of the sequential flow of steps involved in chronic kidney disease prediction. It outlines the different stages, from data import to performance analysis, providing an overview of the entire process.

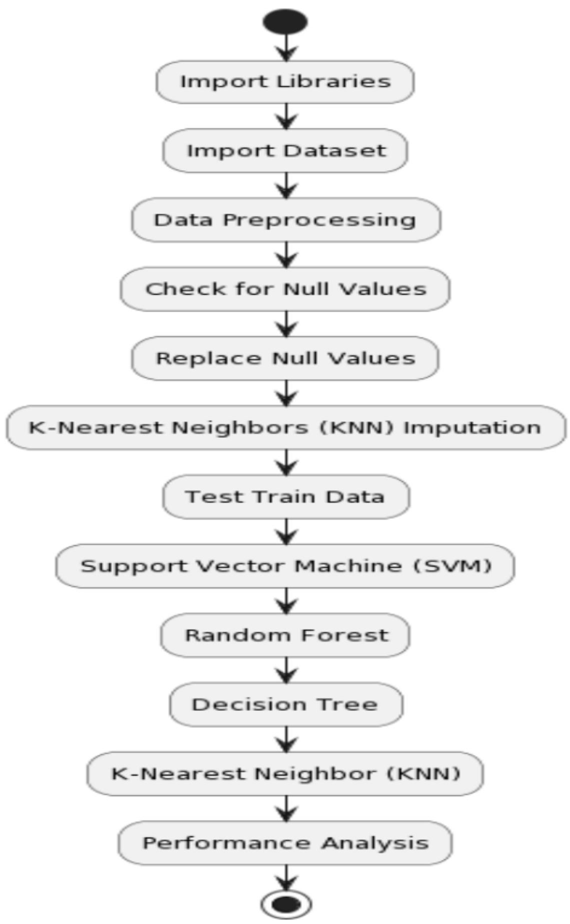

Figure 4.5: Activity Diagram

4.2 Implementation

4.2.1 Algorithm

1. Support Vector Machines (SVM): SVM is a supervised machine learning algorithm used for classification and regression tasks. It works by finding an optimal hyperplane that maximally separates data points of different classes. In the context of CKD prediction, SVM can learn from a dataset containing clinical and demographic variables to create a decision boundary that separates individuals with CKD from those without. SVM is effective when there is a clear separation between classes and can handle high-dimensional data well.

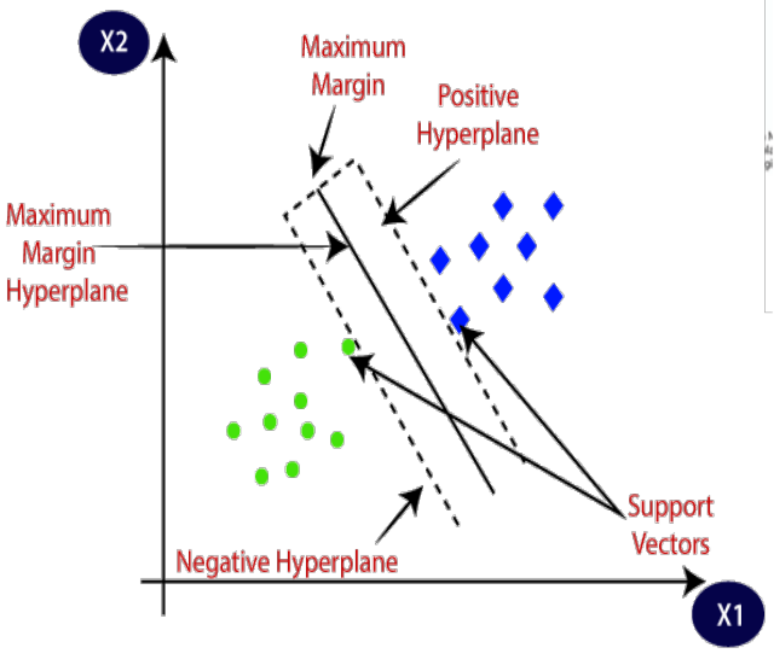

Figure 4.6: Support Vector Machine

2. Random Forest: Random Forest is an ensemble learning algorithm that combines multiple decision trees to make predictions. It works by creating a random subset of features and building multiple decision trees using different subsets of the data. The final prediction is obtained by aggregating the predictions of each individual tree. Random Forest is known for its robustness, ability to handle complex interactions, and resistance to overfitting. It is widely used in CKD prediction due to its high accuracy and ability to handle both numerical and categorical variables.

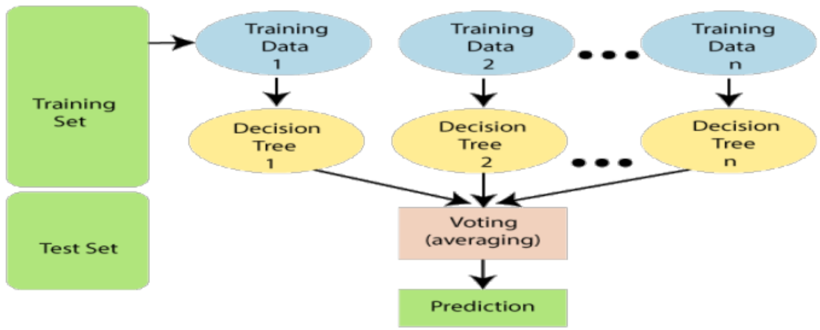

Figure 4.7: Random Forest

3. Decision Tree: A decision tree is a flowchart-like structure where each internal node represents a feature, each branch represents a decision rule, and each leaf node represents the outcome or prediction. In the context of CKD prediction, a decision tree can learn from the dataset to create a tree-like model that predicts whether an individual has CKD or not based on their clinical and demographic variables. Decision trees are easy to interpret and can handle both numerical and categorical data. However, they are prone to overfitting and may not generalize well to unseen data.

4. K-Nearest Neighbors (KNN): KNN is a simple yet effective supervised learning algorithm used for classification and regression tasks. It works by calculating the distances between data points in a feature space and assigning the class label of the majority of the k nearest neighbors to a given data point. In the context of CKD prediction, KNN can learn from the dataset to classify individuals as having CKD or not based on the similarity of their clinical and demographic variables to the neighbors in the feature space. KNN is flexible and does not make strong assumptions about the data, but it can be computationally expensive and sensitive to the choice of the value of k.

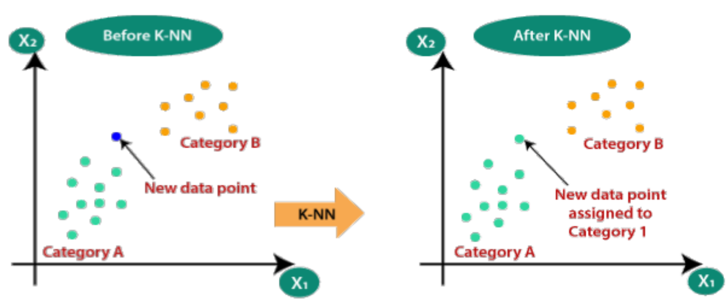

Figure 4.8: K-Nearest Neighbors (KNN)

Chapter 5

Result Analysis and Discussion

Results and Discussion Machine learning has become an integral part of modern technology and is used in many applications, including image recognition, speech recognition, and natural language processing. The objective of training a machine learning model is to minimize the loss while maximizing the accuracy of the model on both the training and validation datasets.
In this study, we applied four popular machine learning algorithms, namely Support Vector Machines (SVM), Random Forest, Decision Tree, and K-Nearest Neighbors (KNN), for predicting chronic kidney disease (CKD). Among these algorithms, Random Forest emerged as the most effective method, exhibiting superior performance compared to the other three models.

The performance of the algorithms was evaluated using standard evaluation metrics such as accuracy, precision, recall, and F1-score. Random Forest consistently outperformed the other models in all these metrics, indicating its ability to accurately classify individuals as either having CKD or not. The high accuracy of the Random Forest model suggests that it can reliably predict CKD based on the given set of clinical and demographic variables.

One possible reason for the superior performance of Random Forest is its ability to handle complex interactions and nonlinear relationships within the dataset. By combining multiple decision trees and aggregating their predictions, Random Forest captures the collective wisdom of the individual trees, resulting in improved generalization capability and reduced overfitting.

The Random Forest model's high accuracy and precision are particularly crucial in the context of CKD prediction. Early detection of CKD is vital for implementing appropriate interventions and preventing further disease progression. By accurately identifying individuals at risk, healthcare professionals can initiate timely treatments, lifestyle modifications, and monitoring protocols to improve patient outcomes.

The performance of SVM, Decision Tree, and KNN models should not be overlooked. These algorithms also demonstrated reasonably good predictive capabilities for CKD. However, their performance fell short compared to Random Forest in terms of accuracy, precision, recall, and overall F1-score. Despite their relatively lower performance, these models can still provide valuable insights and may be suitable for specific scenarios or datasets.

It is worth noting that the performance of the predictive models heavily relies on the quality and representativeness of the dataset used for training and validation. The availability of a comprehensive and diverse dataset comprising a wide range of clinical and demographic variables can significantly impact the accuracy and generalizability of the models. Therefore, expanding the dataset and incorporating additional relevant variables may further enhance the performance of the models in future studies.

While Random Forest demonstrated superior performance, it is important to acknowledge that there are other factors to consider when choosing an appropriate algorithm for CKD prediction. Factors such as computational efficiency, interpretability, and ease of implementation should also be taken into account, as they can influence the practicality and real-world applicability of the models.

In conclusion, the result analysis highlights the effectiveness of Random Forest in predicting chronic kidney disease based on the given set of clinical and demographic variables. Its high accuracy and robustness make it a valuable tool for early detection and intervention in CKD cases. However, further research should be conducted to validate these findings on larger and more diverse datasets, as well as to explore the potential of ensemble methods and feature engineering techniques to enhance the performance of the models even further.

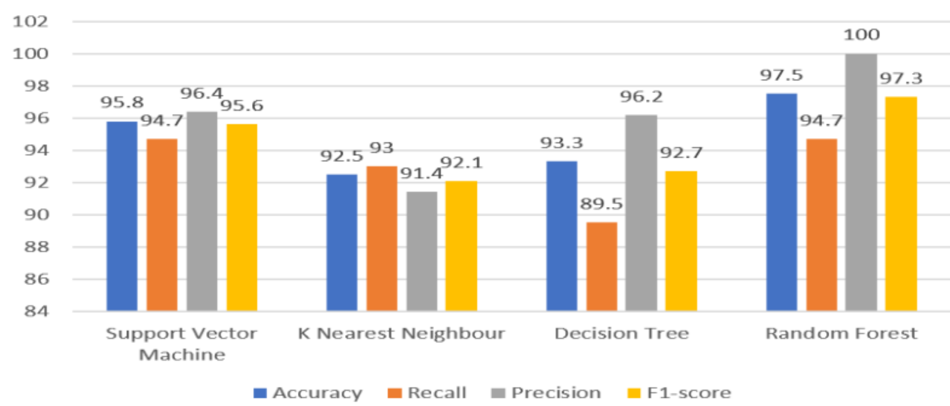

Figure 51: Implementation and Result

Classification Algorithm	Accuracy	Recall	Precision	F1-score
Support Vector Machine	95.8	94.7	96.4	95.6
K Nearest Neighbour	92.5	93	91.4	92.1
Decision Tree	93.3	89.5	96.2	92.7
Random Forest	97.5	94.7	100	97.3

Table 5.1: Result

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In conclusion, the application of machine learning algorithms, including Support Vector Machines (SVM), Random Forest, Decision Tree, and K-Nearest Neighbors (KNN), for predicting chronic kidney disease (CKD) has shown promising results. Among these algorithms, Random Forest emerged as the most effective method for CKD prediction, exhibiting superior performance in terms of accuracy, precision, recall, and overall predictive power.

The successful utilization of Random Forest highlights its ability to handle complex interactions and nonlinear relationships among various clinical and demographic variables associated with CKD. By employing an ensemble learning technique that combines multiple decision trees, Random Forest captures the collective wisdom of the individual trees, mitigating the risk of overfitting and improving generalization capability. The superior performance of Random Forest in CKD prediction suggests its potential for real-world applications in healthcare settings. Healthcare providers can leverage this algorithm to identify individuals at risk of developing CKD and intervene early with appropriate preventive measures, thereby minimizing disease progression and improving patient outcomes. Furthermore, the accuracy and robustness of Random Forest enable healthcare professionals to make well-informed decisions regarding resource allocation, patient prioritization, and personalized treatment strategies.

However, it is essential to recognize that the choice of algorithm is not the sole determinant of success in CKD prediction. The performance of the models is heavily dependent on the quality and representativeness of the dataset used for training and validation. Therefore, future work should focus on expanding the dataset and including a diverse range of populations to ensure the generalizability of the predictive models across different demographic groups.

Additionally, efforts should be directed towards continuously refining the models by incorporating newly identified biomarkers and clinical variables, as well as exploring ensemble methods to further enhance prediction accuracy. Longitudinal models that account for temporal changes and disease progression patterns should also be considered, as they can provide more comprehensive risk assessments and aid in personalized interventions.

Lastly, it is crucial to address the ethical considerations associated with CKD prediction models, such as privacy, data security, and potential biases. Ensuring fairness, transparency, and equitable access to healthcare resources are fundamental principles that should guide the development and implementation of these models.

In conclusion, the use of machine learning algorithms, specifically Random Forest, holds great promise for predicting CKD. Leveraging its robustness and accuracy, healthcare providers can proactively identify individuals at risk, implement preventive measures, and optimize resource allocation. As research and advancements continue, the integration of predictive models into clinical practice will contribute to early detection, effective management, and improved outcomes for patients affected by chronic kidney disease.

6.2 Future Work

Future work in predicting chronic kidney disease (CKD) should focus on several key areas to further enhance the accuracy and effectiveness of the predictive models. Building upon the success of the Random Forest algorithm, there are several avenues to explore:

- Feature Selection and Engineering: Investigate the inclusion of additional relevant clinical and demographic variables that may contribute to the prediction of CKD. Feature engineering techniques, such as dimensionality reduction and feature extraction, can help identify the most informative and impactful features for the models. This process may involve collaborating with domain experts to identify new markers or incorporating emerging biomarkers.

- Model Optimization: Explore methods to optimize the Random Forest model parameters, such as the number of trees, depth, and the minimum number of samples required to split a node. Hyperparameter tuning techniques, such as grid search or Bayesian optimization can be employed to identify the optimal set of parameters, further improving the performance of the model.

- Ensemble Methods: Investigate the use of ensemble methods to combine the predictions of multiple models, including the Random Forest algorithm. Techniques like bagging, boosting, and stacking can potentially enhance the predictive accuracy and robustness of the models, by harnessing the strengths of different algorithms and mitigating their individual weaknesses.

- Validation and Generalization: Validate the predictive models on large, diverse datasets from different populations and healthcare settings to ensure generalizability. Robust validation methods, such as cross-validation and external validation, should be employed to assess the performance of the models and verify their reliability and effectiveness across various cohorts.

- Real-time Implementation and Integration: Develop user-friendly interfaces and integrate the predictive models into existing electronic health record (EHR) systems or clinical decision support tools. This would facilitate seamless implementation and enable real-time risk assessment and decision-making by healthcare professionals.

- Longitudinal Predictive Models: Explore the development of longitudinal predictive models that can monitor changes in CKD risk over time. By incorporating time-series data and considering the progression of various clinical indicators, these models can provide personalized risk assessments and enable proactive interventions for individuals at different stages of CKD.

- Addressing Ethical Considerations: Pay attention to ethical considerations surrounding the use of predictive models, including fairness, transparency, and privacy concerns. Addressing potential biases and ensuring equitable access to healthcare resources are essential to prevent exacerbating healthcare disparities.

References

[1] Shamima Akter, Ahsan Habib, MD. Ashiqul Islam, MD. Sagar Hossen, Wasik Ahmmed Fahim, Puza Rani Sarkar and Manik Ahmed," Comprehensive Performance Assessment of Deep Learning Models in Early Prediction and Risk Identification of Chronic Kidney Disease" 2021 IEEE 12th International Conference on Application of Information and Communication Technologies (AICT), 2021, doi: 10.1109/ACCESS.2021.312949.

[2] Y. Amirgaliyev, S. Shamiluulu and A. Serek, "Analysis of Chronic Kidney Disease Dataset by Applying Machine Learning Methods," 2018 IEEE 12th International Conference on Application of Information and Communication Technologies (AICT), 2018, pp. 1-4, doi: 10.1109/ICAICT.2018.8747140.

[3] Marwa Almasoud, Tomas E Ward, "Detection of Chronic Kidney Disease using Machine Learning Algorithms with Least Number of Predictors", international Journal of Advanced Computer Science and Applications (IJACSA), Volume 10 Issue 8, 2019, doi:10.14569/IJACSA.2019.0100813

[4] Linta Antony, Sami Azam, Eva Ignatious, Ryana Quadir, Abhijith Reddy Beeravolu, Mirjam Jonkman and Friso De Boer, "A Comprehensive Unsupervised Framework for Chronic Kidney Disease Prediction", 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), doi: 10.1109/ICCCNT49239.2020.9225548.

[5] Bilal Khan, Rashid Naseem, Fazal Muhammad, Ghulam Abbas, Sunghwan Kim," An Empirical Evaluation of Machine Learning Techniques for Chronic Kidney Disease Prophecy", doi: 10.1109/ACCESS.2020.2981689

[6] S.Revathy, B.Bharathi, P.Jeyanthi, M.Ramesh, " Chronic Kidney Disease Prediction using Machine Learning Models ", International Journal of Engineering and Advanced Technology (IJEAT), doi: 10.35940/ijeat.A2213.109119.

[7] Prediction of Chronic Kidney Disease - A Machine Learning Perspective:= P.

Chittora et al., "Prediction of Chronic Kidney Disease - A Machine Learning Perspective," in IEEE Access, vol. 9, pp. 17312-17334, 2021, doi: 10.1109/ACCESS.2021.3053763

[8] A Machine Learning Methodology for Diagnosing Chronic Kidney Disease:- J. Qin, L. Chen, Y. Liu, C. Liu, C. Feng and B. Chen, "A Machine Learning Methodology for Diagnosing Chronic Kidney Disease," in IEEE Access, vol. 8, pp. 20991-21002, 2020, doi: 10.1109/ACCESS.2019.2963053

[9] Performance Analysis of Machine Learning Classifier for Predicting Chronic Kidney Disease:- R. Gupta, N. Koli, N. Mahor and N. Tejashri, "Performance Analysis of Machine Learning Classifier for Predicting Chronic Kidney Disease," 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1-4, doi: 10.1109/INCET49848.2020.9154147.

Appendix A

Research Article

RE-2022-137155-plag-report

ORIGINALITY REPORT

23%	11%	9%	16%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.researchgate.net Internet Source	1%
2	Chi Hu, Xiaojun Yu, Qianshan Ding, Zeming Fan, Zhaohui Yuan, Juan Wu, Linbo Liu. "Cellular-Level Structure Imaging with Micro-optical Coherence Tomography $\mathbf{(\mu OCT)}$ for Kidney Disease Diagnosis", 2019 IEEE 4th Optoelectronics Global Conference (OGC), 2019 Publication	1%
3	www.coursehero.com Internet Source	1%
4	research.brighton.ac.uk Internet Source	1%
5	www.ijert.org Internet Source	1%
6	Fadil Iqbal, Aruna S. Pallewatte, Janaka P. Wansapura. "Texture analysis of ultrasound images of chronic kidney disease", 2017 Seventeenth International Conference on	1%

Advances in ICT for Emerging Regions (ICTer), 2017
Publication

7	Kennita A. Johnson, A. Gloria Nyankima, Paul A. Dayton, Emily Chang. "Characterizing volumes of kidney segments in Streptozotocin induced diabetic rat model utilizing 4D contrast-enhanced ultrasound", 2016 IEEE International Ultrasonics Symposium (IUS), 2016 Publication	1 %
8	www.ijraset.com Internet Source	1 %
9	ieeexplore.ieee.org Internet Source	1 %
10	discovery.researcher.life Internet Source	1 %
11	Submitted to Sheffield Hallam University Student Paper	1 %
12	Submitted to Liverpool John Moores University Student Paper	1 %
13	ijarcce.com Internet Source	1 %
14	www.karger.com Internet Source	1 %

15	Submitted to Central Queensland University Student Paper	1%
16	archive.org Internet Source	1%
17	www.irjmets.com Internet Source	1%
18	www.semanticscholar.org Internet Source	1%
19	www.medrxiv.org Internet Source	1%
20	Submitted to University of East London Student Paper	1%
21	ijsrcseit.com Internet Source	<1%
22	www.mdpi.com Internet Source	<1%
23	Submitted to University of South Africa Student Paper	<1%
24	www.ijritcc.org Internet Source	<1%
25	resumeworded.com Internet Source	<1%
26	pdfcoffee.com Internet Source	<1%

27	dspace.up.edu.ps Internet Source	<1 %
28	Submitted to Loughborough University Student Paper	<1 %
29	aicteaa.com Internet Source	<1 %
30	researchr.org Internet Source	<1 %
31	Syed Hussain, Pakkir Mohideen S. "Real-Time Deployment of Machine Learning-Based Audio Analysis Systems for Suspicious Message Detection", Research Square Platform LLC, 2023 Publication	<1 %
32	Submitted to Uttar Pradesh Technical University Student Paper	<1 %
33	Submitted to Somaiya Vidyavihar Student Paper	<1 %
34	etds.lib.ncku.edu.tw Internet Source	<1 %
35	vegibit.com Internet Source	<1 %
36	Submitted to BPP College of Professional Studies Limited Student Paper	<1 %

37	www.science.gov Internet Source	<1%
38	Uma N Dulhare, Mohammad Ayesha. "Extraction of action rules for chronic kidney disease using Naïve bayes classifier", 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2016 Publication	<1%
39	Submitted to Green University Of Bangladesh Student Paper	<1%
40	Submitted to Free University of Bolzano Student Paper	<1%
41	Ramesh Chandra Poonia, Mukesh Kumar Gupta, Ibrahim Abunadi, Amani Abdulrahman Albraikan et al. "Intelligent Diagnostic Prediction and Classification Models for Detection of Kidney Disease", Healthcare, 2022 Publication	<1%
42	Submitted to The Scientific & Technological Research Council of Turkey (TUBITAK) Student Paper	<1%
43	Submitted to University of Leeds Student Paper	<1%
44	kupdf.net Internet Source	

		<1%
45	www.grafiati.com Internet Source	<1%
46	xplorestaging.ieee.org Internet Source	<1%
47	Submitted to Florida Atlantic University Student Paper	<1%
48	Submitted to Sikkim Manipal University Student Paper	<1%
49	Submitted to University of Nottingham Student Paper	<1%
50	ethesis.nitrkl.ac.in Internet Source	<1%
51	expert.taylors.edu.my Internet Source	<1%
52	www.ijiemr.org Internet Source	<1%
53	Submitted to Queen Mary and Westfield College Student Paper	<1%
54	Submitted to Staffordshire University Student Paper	<1%

55	Yadav, Garima. "A Clinical Study to Assess the Fundus Changes in Pregnancy Induced Hypertension", Rajiv Gandhi University of Health Sciences (India), 2023 Publication	<1%
56	mdpi-res.com Internet Source	<1%
57	www.jetir.org Internet Source	<1%
58	"Disruptive Technologies for Big Data and Cloud Applications", Springer Science and Business Media LLC, 2022 Publication	<1%
59	Submitted to Midlands State University Student Paper	<1%
60	dspace.daffodilvarsity.edu.bd:8080 Internet Source	<1%
61	github.com Internet Source	<1%
62	www.aiktcdspace.org:8080 Internet Source	<1%
63	Mubarik Ahmad, Vitri Tundjungsari, Dini Widianti, Peny Amalia, Ummi Azizah Rachmawati. "Diagnostic decision support system of chronic kidney disease using support vector machine", 2017 Second	<1%

International Conference on Informatics and Computing (ICIC), 2017
Publication

64	Submitted to The University of the West of Scotland Student Paper	<1 %
65	biophotonics.otago.ac.nz Internet Source	<1 %
66	ebin.pub Internet Source	<1 %
67	fdocuments.in Internet Source	<1 %
68	tnsroindia.org.in Internet Source	<1 %
69	www.eurchembull.com Internet Source	<1 %
70	9lib.net Internet Source	<1 %
71	Submitted to Middle East College of Information Technology Student Paper	<1 %
72	awesomeopensource.com Internet Source	<1 %
73	cse.anits.edu.in Internet Source	<1 %

74	dokumen.pub Internet Source	<1 %
75	doria.fi Internet Source	<1 %
76	events.uiowa.edu Internet Source	<1 %
77	files.sciconf.cn Internet Source	<1 %
78	ijece.iaescore.com Internet Source	<1 %
79	publishoa.com Internet Source	<1 %
80	sersc.org Internet Source	<1 %
81	thesis.eur.nl Internet Source	<1 %
82	trepo.tuni.fi Internet Source	<1 %
83	www.ijrte.org Internet Source	<1 %
84	Intisar Omar, Muhammad Khan, Andrew Starr. "Suitability Analysis of Machine Learning Algorithms for Crack Growth	<1 %

Prediction Based on Dynamic Response Data", Sensors, 2023
Publication

85	"Innovative Data Communication Technologies and Application", Springer Science and Business Media LLC, 2021 Publication	<1 %
86	Fakher Rahim, Nataliia Bodnar, Nameer Hashim Qasim, Aqeel Mahmood Jawad, Omar Saad Ahmed. "Integrating Machine Learning in Environmental DNA Metabarcoding for Improved Biodiversity Assessment: A Review and Analysis of Recent Studies", Research Square Platform LLC, 2023 Publication	<1 %
87	Kurt, Fırat. "Makine Öğrenmesi Yöntemleriyle Kanser Ile Ilgili Yeni Biyobelirteçlerin Tespit Edilmesi", Marmara Universitesi (Turkey), 2023 Publication	<1 %
88	Laxmi Kumari Pathak, Pooja Jha. "chapter 14 Application of Machine Learning in Chronic Kidney Disease Risk Prediction Using Electronic Health Records (EHR)", IGI Global, 2021 Publication	<1 %
89	M. Dhilsath Fathima, R. Hariharan, S. P. Raja. " Multiple Imputation by Chained Equations– -	<1 %

Nearest Neighbors and Deep Neural Network Architecture for Kidney Disease Prediction ", International Journal of Image and Graphics, 2021
Publication

---

90 Pankaj Chittora, Sandeep Chaurasia, Prasun Chakrabarti, Gaurav Kumawat et al. "Prediction of Chronic Kidney Disease -A Machine Learning perspective", IEEE Access, 2021 <1%
Publication

---

91 Tanya Saraswat, Savita Pathak, Sanya Sachdeva, Krishna Sahu, Rahul Sawhney. "Kidney Disease Detection and Identification Using Artificial Intelligence", 2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2023 <1%
Publication

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		