VISVESVARAYA TECHNOLOGICAL UNIVERSITY

Jnana Sangama, Belagavi-590018, Karnataka

A PROJECT REPORT

ON

"Prediction of Chronic Kidney Disease Using Machine Learning"

Submitted in partial fulfillment of requirements for the award of degree of Bachelor of Engineering

in

COMPUTER SCIENCE & ENGINEERING

Submitted By:

1. ARCHANA S MALAKHED	1MJ18CS022
2. DEEPTHI H P	1MJ18CS035
3. DHARANI K	1MJ18CS037
4. G LAKSHMI KEERTHI	1MJ18CS048

Under the Guidance of

Mrs. S. SUGUNA

Assistant Professor,

Department of Computer Science &Engineering

COLLEGE OF ENGINEERING
Since 1982
Engineering A Better Tomorrow

An Autonomous Institute

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
MVJ COLLEGE OF ENGINEERING
BANGALORE-67

ACADEMIC YEAR-2021-22

An Autonomous Institute

MVJ COLLEGE OF ENGINEERING

Near ITPB, Whitefield, Bangalore-67

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the project work, entitled "**PREDICTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING"** is a bonafide work carried out by

1. ARCHANA S MALAKHED	1MJ18CS022
2. DEEPTHI H P	1MJ18CS035
3. DHARANI K	1MJ18CS037
4. G LAKSHMI KEERTHI	1MJ18CS048

in partial fulfillment for the award of degree of Bachelor of Engineering in Computer Science & Engineering of the Visvesvaraya Technological University, Belagavi during the academic year 2021-22. It is certified that all the corrections/suggestions indicated for Internal Assessment have been incorporated in the Report. The project report has been approved as it satisfies the academic requirements.

Signature of the Internal Guide	**Signature of the HOD**	**Signature of the Principal**

Name of examiners: Signature with date

1.
2.

An Autonomous Institute
MVJ COLLEGE OF ENGINEERING
Whitefield, Near ITPB, Bangalore-67

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DECLARATION

We, hereby declare that the entire work titled "**PREDICTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING"** embodied in this project report has been carried out by us during the 8th semester of BE degree at MVJCE, Bangalore under the esteemed guidance of **Mrs. S. Suguna**, Assistant. Prof, Dept. of CSE, MVJCE affiliated to Visvesvaraya Technological University, Belagavi. The work embodied in this dissertation work is original and it has not been submitted in part or full for any other degree in any University.

1. ARCHANA S MALAKHED	1MJ18CS022	_____
2. DEEPTHI H P	1MJ18CS035	_____
3. DHARANI K	1MJ18CS037	_____
4. G LAKSHMI KEERTHI	1MJ18CS048	_____

Place:

Date:

ABSTRACT

One of the biggest health problems in the world is chronic kidney disease (CKD), which can lead to renal failure, cardiovascular disease (CVD), and early death. For patients who progress to end-stage kidney disease, insufficient availability to kidney replacement therapy is a problem in many regions of the world. According to research, 1 in 10 people may have CKD, however it is more prevalent in adults than children, where it affects 1 in 50 people and 1 in 2 persons over the age of 75. Millions of people worldwide die each year from CKD, which affects 10% of the population and typically has no symptoms due to aberrant test findings. Data classification can be used to identify CKD, but the system might not be able to accurately estimate the proportion of CKD patients. The innovative aspect of this study is the method used to diagnose chronic renal disease. The goal of this project is to compare the accuracy-based performance of different machine learning algorithms employing Liner Classifier, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Naive Bayes, Random Forest, and XGBoost.

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany a successful completion of any task would be incomplete without the mention of people who made it possible, success is the epitome of hard work and perseverance, but steadfast of all is encouraging guidance.

So, with gratitude we acknowledge all those whose guidance and encouragement served as beacon of light and crowned our effort with success.

We are thankful to our **Management** for being a constant inspiration and providing all the facilities that needed throughout the project work.

We would like to thank our **Principal Dr. P. Mahabaleswarappa,** and for his constant guidance, advice and encouragement to complete the project work.

We would like to thank our **Vice principal Dr. Brindha M**, for her incessant encouragement & all the help during the project work.

We would also like to express our sincere gratitude to our **Dr. M A Lourdu Antony Raj** Controller of Examinations, MVJCE, Bengaluru for his persistent guidance.

We consider it a privilege and honor to express our sincere gratitude to our beloved **HOD Mrs. R. Tamilarasi** for her constant encouragement and all the support provided during this project work.

We convey our sincere thanks to our guide **Mrs. S. Suguna, Assistant Professor, Dept. of CSE** for her valuable guidance throughout the tenure of this project work, and whose support and encouragement made this work possible.

It's also a great pleasure to express our deepest gratitude to all our faculty members of our department for their cooperation and constructive criticism offered, which helped us a lot during our project work.

Finally, we would like to thank all our family members and friends whose encouragement and support was invaluable.

Thank you

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Project Overview

Chronic kidney disease (CKD) is a world's major health issue, with adverse outcomes of kidney failure, cardiovascular disease (CVD), and premature death. In many parts of the world, poor access to kidney replacement therapy is an issue for patients who advance to end-stage kidney disease. Research suggests that 1 in 10 of the population mayhave CKD, but it is common in adults, being present in 1 in 50 people, for those aged over75, CKD is present in 1 out of 2 people. Of the world's population, 10% are affected by CKD and millions die each year, as in most cases it doesn't cause any symptoms due to theabnormality in the test results. CKD can be detected by classifying the data but the systemmay fail to give an accurate percentage of the people having CKD. The novelty of this study lies in the diagnosis system to detect chronic kidney disease. This project's aim is todifferentiate the performance of various machine learning algorithms that are primarily based on their accuracy by using Liner Classifier, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Naïve Bayes, Random Forest, and XG Boost.

1.2 Purpose of Project

The sole purpose of this project is to detect CKD accurately based on features of the CKDdataset. In this KNN, Linear classifier, Random Forest, SVM, XG Boost, and Naïve BayesClassifier algorithms are used. For each classifier, results were computed based on the accuracy, precision, recall, and F1 score.

1.3 Scope of Project

More than 1 in 7 adults in the United States may have chronic kidney disease (CKD), withhigher rates in low-income and racial/ethnic minority groups. And most people with CKDdon't know they have it. Healthy People 2030 focuses on preventing, diagnosing, and treating CKD. People with CKD are more likely to have heart disease and stroke and to dieearly. Managing risk factors like diabetes and high blood pressure can help prevent or delayCKD. Strategies to make sure more people with CKD are diagnosed early can help peopleget the treatment they need. Recommended tests can help identify people with CKD to make sure they get treatments and education that may help prevent or delay kidney failure

and end-stage kidney disease (ESKD). In addition, strategies to make sure more people with ESKD get kidney transplants can increase survival rates and improve quality of life.

1.4 Problem Statement

Chronic kidney disease (CKD) is the presence of lasting functional abnormalities in the renal structure and function that can worsen over time. Kidneys are vital for the whole bodyto work properly. Therefore, when a person has CKD, they may also have problems with the functioning of the rest of their body. Some of the common complications of CKD include anemia, bone disease, heart disease, high potassium, high calcium, and fluid accumulation.

1.5 Existing System

Machine learning refers to a computer program, which calculates and deduces the information related to the task and obtains the characteristics of the corresponding pattern.This technology can achieve accurate and economical diagnoses of diseases; hence, it mightbe a promising method for diagnosing CKD. It has become a new kind of medical tool withthe development of information technology and has a broad application prospect because of the rapid development of electronic health records. In the medical field, machine learninghas already been used to detect human body status, analyze the relevant factors of the disease and diagnose various diseases.

For example, the models built by machine learning algorithms were used to diagnose heartdisease, diabetes and retinopathy, acute kidney injury, cancer, and other diseases. In these models, algorithms based on regression, tree, probability, decision surface, and neural network were often effective. In the field of CKD diagnosis, Hodneland et al. utilized imageregistration to detect renal morphologic changes. Vasquez-Morales et al. established a classifier based on the neural network using large-scale CKD data, and the accuracy of themodel on their test data was 95%. In addition, most of the previous studies utilized the CKD data set that was obtained from the UCI machine learning repository. Chen et al. usedk-nearest neighbor (KNN), support vector machine (SVM), and soft independent modellingof class analogy to diagnose CKD, KNN and SVM achieved the highest accuracy of 95.2%. In addition, they used fuzzy rule-building expert system, fuzzy optimal associative memoryand partial least squares discriminant analysis to diagnose CKD, and the range of accuracyin those models was 94.8%.

Current System do not used many algorithms for prediction purpose. Dataset used by current system is not having more parameter. Algorithms which are used by current systemare not giving highest accuracy.

1.5.1 Disadvantages of Existing System

In the above model, the mean imputation is used to fill in the missing values and it dependson the diagnostic categories of the samples. As a result, existing method could not be usedwhen the diagnostic results of the samples are unknown. In reality, patients might miss some measurements for various reasons before diagnosing. In addition, for missing valuesin categorical variables, data obtained using mean imputation might have a large deviationfrom the actual values. For example, for variables with only two categories, we set the categories to 0 and 1, but the mean of the variables might be between 0 and 1. Polat et al. developed an SVM based on feature selection technology, the proposed models reduced the computational cost through feature selection, and the range of accuracy in those modelswas from 94.75%-95.5%. J. Aljaaf et al. used novel multiple imputation to fill in the missing values, and then MLP neural network

(MLP) achieved an accuracy of 96.1%. The studies of, focus mainly on the establishmentof models and achieve an ideal result. However, a complete process of filling in the missingvalues is not described in detail, and no feature selection technology is used to select predictors as well. Almansour et al. used SVM and neural network to diagnose CKD, and the accuracy of the models was 94.75% and 95.75%, respectively. In the models establishedby Gunarathne et al., zero was used to fill out the missing values and decision forest achieved the best performance with the accuracy was 94.1%. To summarize the previous CKD diagnostic models, we find that most of them suffering from either the method used to impute missing values has a limited application range or relatively low accuracy.

Therefore, in this work, we propose a methodology to extend application range of the CKD diagnostic models.

1.6 Proposed System

This article objects to predict Chronic Kidney Disease based on full features and importantfeatures of CKD dataset. For feature selection different technique has been applied: Featureselection. In this perception, six classifiers algorithm were applied viz. Linear Classifier, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Naïve Bayes, Random Forest, and XGBoost. For each classifier, the results were computed based on full features, Accuracy achieved by different classifier are: Linear Classifier is giving accuracy about 96%, KNN is achieving accuracy of 96%, SVM is giving accuracy of 95%, Naïve Bayes classifier is giving 96%, Random Forest and XGBoost is giving accuracy of 98%. Result is concluded by considering accuracy, precision, recall and F1-score. Using these machinelearning algorithms, we are proposing this CKD system for classification.

1.6.1 Objectives

Chronic Kidney Disease is one of the most critical illnesses nowadays and proper diagnosis is required as soon as possible. Machine learning technique has become reliable for medicaltreatment. With the help of a machine learning classifier algorithms, the doctor can detect the disease on time. Chronic Kidney Disease dataset has been taken from the Kaggle repository. Six classifier algorithms have been applied in this research such as Linear Classifier, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Naïve Bayes, Random Forest, and XGBoost. Linear Classifier is giving accuracy about 96%, KNN is achieving accuracy of 93%, SVM is giving accuracy of 97%, Naïve Bayes classifier is giving 96%, Random Forest and XGBoost is giving accuracy of 96%. Result is concluded by considering accuracy, precision, recall and F1-score. Among these algorithms, XGBoostis giving more accuracy of 98%.

Summary

This chapter gives an overview of project, its scope, existing system and its disadvantages, proposed system and objectives of proposed system.

CHAPTER 2

LITERATURE STUDY

2.1 Overview

Literature study is carried out to get all the related information of current project, which isused to get an idea for the enhancement as well as changes that can be made to improve existing approaches. A literature study is done on various machine learning techniques. Following section describes about all the related papers which is used in current project.

2.2 Related Papers

PAPER 1: **Kidney stone recognition and extraction using Directional Emboss & SVMfrom Computes Tomography Images**

"Akanksha Soni al" The intention of this paper is to develop an automatic stone detection system from the Computed Tomography (CT) Image. The primary aspire is to inspect the CT scan to recognize whether the kidneys are influenced by a stone. The algorithm used isthe SVM (Support Vector Machine) to classify the stone. SVM is a linear model for classification and regression problems [1]. It can solve linear and non-linear problems andwork well for many problems. This algorithm creates a line or hyperplane which separatesthe data into classes.

SVM is used for classifying two different classes of data in two particular shells and construct a hyperplane that can be valuable for regression. Emboss is a computer graphicstechnique in which each pixel of an image is replaced either by a highlight or a shadow, depending on light/dark boundaries on the original image. Low contract areas are replacedby a gray background.

The proposed system helps to detect the stone in the kidney but fails to identify multiple stone in kidney. In the future scope it is suggest to develop the system which could detect multiple number of stones in the kidney with more accuracy [1]. Automatic identificationis one of the trending in medical science which is helpful for saving human health and wealth as well.

PAPER 2: Kidney Stone Detection Using Neural Networks

"M.Akshaya al" Kidney may fail functioning due to the stone formation in the kidney. Thatis highly impractical to handle the large amount of data to detect manually. Kidneys removewastes and extra fluid from our body. Kidneys also removes acid that is produced by the cells of our body and maintain a healthy balance of water, salts and minerals such as sodium, calcium, phosphorous and potassium in our body [2]. Kidney may fail functioningdue to the stone formation in the kidney. That is highly impractical to handle the large amount of data to detect manually. By using AI based methods along with neural networksand feature extraction has shown the great potential using back propagation network was applied to detect the stone in the kidney.

The feature extraction is done by using principal component analysis and the image classification is done using back propagation network. Back propagation gives a perfect classification with greater accuracy.

The algorithm used is the Back Propagation Network (BPN) [2]. BPN is used to detect stone in the kidney. BPN gives a quick and precise classification and one of the better toolsfor identifying the stone in the kidney. This methos also presents a segmentation method using FCM (Fuzzy C-Mean) clustering algorithm.

PAPER 3: Explainable Prediction of Chronic Renal Disease in the Colombian Population Using Neural Networks and Case-Based Reasoning

Gabriel R. Vasquez- Morales al" This paper presents a neural- network primarily based classifier that helps in predicting whether or not someone is in danger of developing Chronic excretory organ Disease (CKD). Nowadays computer science plays a very important role in health sector. Machine Learning Algorithms are wide utilized in malady prediction. Some Algorithms like Support Vector Machine (SVM), Random Forest, NeuralNetworks are accustomed predict and classify patients' polygenic disorder, heart condition,cancer et al. Neural Networks have currently been used as support tools for early malady detection. This paper presents a neural- network primarily based classifier that helps in predicting whether or not someone is in danger of developing Chronic excretory organ Disease (CKD). The model is trained with demographic knowledge and treatment info of 2 population teams.

Current analysis proposes the employment of dual system. Twin system refers Neural Network (NN)- Case primarily based Reasoning (CBR) that is applied and valid for the

reason of CKD predictions. Neural Networks are accustomed notice CKD and alternative effects associated with urogenital apparatus.

PAPER 4: A Comprehensive Unsupervised Frame work for Chronic Kidney Disease Prediction

"L. Antony et al" Early detection helps to provide proper treatment and care to patients. Inmany hospitals there is a shortage of nephrologists so patients have to wait for a longer to get diagonize. CKD decreases the kidney function over time as indicated by glomerular filtration rate (GMR) and markers of kidney damage. People with CKD are dying at early age. It is crucial to detect CKD at an early stage. Early detection helps to provide proper treatment and care to patients. In many hospitals there is a shortage of nephrologists so patients have to wait for a longer to get diagonize. Thus, with the help of AI technology this system classifies patients into CKD or Non-CKD which will help doctors to treat multiple patients faster.

Dataset is online data repository of the University of California Irvine (UCI) and contains data of 400 patients. It consists of 1 class attribute and 24 clinical attributes. The dataset consists of 250 CKD and 150 non-CKD cases. Dataset includes missing data as well.

Unsupervised learning extracts patterns from CKD related clinical data. Various feature selection mechanisms related to filter methods wrapper methods, embedded methods and unsupervised methods are implemented to identify important features. algorithms such as K-Means Clustering, Isolation Forest, DB-scan and Auto encoder are implemented on selected features. Evaluation metrics are generated to compare with existing models. In classifying CKD with Non-CKD, K-means clustering achieved the accuracy of 94%**.**

PAPER 5: Chronic Kidney disease prediction using Machine Learning

"Pankaj chittora al" Chronic kidney disease (CKD) means kidneys are damaged and are not filtering the blood way it should. if the person has suffered from the CKD means that waste is collected in the body. This disease is chronic because of the damage gradually over a long period. The causes of the CKD include diabetes, high blood pressure, heart disease, it also depends on age and gender. CKD does not give any symptoms till kidney isdamaged badly. There are few diagnostic tests to check the condition of CKD: 1. estimatedglomerular filtration rate (eGFR) 2. Urine test 3. Blood pressure. Algorithm used are FilterMethod, Wrapper Method, and Embedded Method. The primary role of kidney is to filter the extra water and waste from blood to produce urine. Chronic kidney disease (CKD)

means kidneys are damaged and are not filtering the blood way it should. If the person hassuffered from the CKD means that waste is collected in the body. This disease is chronicbecause of the damage gradually over a long period. The causes of the CKD include diabetes, high blood pressure, heart disease, it also depends on age and gender. CKD doesnot give any symptoms till kidney is damaged badly [11]. There are few diagnostic tests tocheck the condition of CKD: 1. estimated glomerular filtration rate (eGFR) 2. Urine test 3.Blood pressure.

A. eGFR

eGFR value shows that how your kidney cleaning the blood. If your eGFR value is greaterthan 90, that means the kidney is normal. If eGFR value is less than 60, that means you have CKD.

B. Urine Test

The doctor also asks for urine test for kidney functionality because kidneys make urine. Ifthe urine contains blood and protein, that means your kidney is not working properly.

C. Blood Pressure

Blood pressure range shows how your heart is pumping blood. If eGFR value reaches lessthan 15, that means the patient has end-stage kidney disease. At this point, there are only available treatments: (i) dialysis and (ii) kidney transplant. If dialysis is not possible, the doctor has only one solution, i.e., kidney transplantation.

Therefore, it is necessary to early recognition, monitoring and handling of the disease. sinceit does not show the symptoms at early stage it is essential to predict the striding of CKD with appropriate accuracy. The machine learning can also find data analysis and pattern detection. A variety of health dataset is present so machine learning algorithms are best fitto improve the accuracy of diagnosis prediction. As healthcare electronic dataset grows rapidly, machine learning algorithms are becoming more common in healthcare.

PAPER 6: An image pre-processing method for kidney stone segmentation in CT scanimages

“Nilar Thein al” Automatic Kidney stone detection segmentation from CT images is challenging. The purpose of this system is to develop reader independent pre-processing

algorithm to assist the segmentation process [15]. Three algorithms based on the intensity,size and location are applied for unwanted regions removing such as soft-organ removing,bony skeleton removing and bed-mat removing. The digitalized transverse abdomen CT scans images from 30 patients with kidney stone cases were included in statistical analysisand validation.

The used algorithm takes image as an input recognizes soft-organ removing and bony skeleton removing and bed-mat removing and outputs the stone detected image as output.

The proposed system method for kidney stone segmentation will provide good support in detection of kidney stone. Intensity-based thresholding can produce the output image without any soft regions. Area-based thresholding can reduce many false positive and bedmat region. By using thresholding method based on the prior knowledge of the image, thepropose system is simple and easy to understand for the segmentation.

2.3 Literature Study

The most of the literature survey focuses on the techniques to detect chronic kidney diseaseusing machine learning algorithms. However, they are not giving more accuracy. Literaturesurvey made us think more logically in order to get more accuracy.

Summary

This chapter discuss about the literature study for the proposed system, related paper whichgives idea of enhancements to the existing work. The related paper gives information aboutthe existing systems, its advantages and disadvantages.

CHAPTER 3

SYSTEM REQUIREMENT

The system requirements specify features, components and behaviour of system which is to be developed. The following sections describe about functional, non-functional, performance related, features and behaviour of the solution. This includes the detailed description of the solution to be developed.

3.1 Resource Requirement

Jupyter Notebook is a non-profit organization created to "develop open-source software, open-standards, and services for interactive computing across dozens of programming languages". Spun-off from IPython in 2014 by Fernando Pérez, Project Jupyter supports execution environments in several dozen languages. Project Jupyter's name is a reference to the three core programming languages supported by Jupyter, which are Julia, Python and R, and also an homage to Galileo's notebooks recording the discovery of the moons of Jupiter. Project Jupyter has developed and supported the interactive computing products Jupyter Notebook, Jupyter Hub, and Jupyter Lab, the next-generation version of Jupyter Notebook. Jupyter Notebook (formerly IPython Notebooks) is a web-based interactive computational environment for creating Jupyter notebook documents. The "notebook" term can colloquially make reference to many different entities, mainly the Jupyter web application, Jupyter Python web server, or Jupyter document format depending on context. A Jupyter Notebook document is a JSON document, following a versioned schema, and containing an ordered list of input/output cells which can contain code, text (using Markdown), mathematics, plots and rich media, usually ending with the ".ipynb" extension.

3.2 Hardware Requirement

The following is the hardware requirements of the system for the proposed system:

Hardware Requirements:

Processor	: Intel(R) Core(TM) i3-8145U CPU @ 2.10GHz 2.30 GHzRAM
	: 256 MB (min)
Storage	: 20 GB
Standard Devices	: Keyboard, Monitor and Mouse

3.3 Software Requirement

The following is the software requirements of the system for the proposed system:

- OS **:** Windows
- Language **:** Python
- IDE/tool **:** Jupyter Notebook
- Dataset **:** https://www.kaggle.com/datasets/mansoordaku/ckdisease

CHAPTER 4

SYSTEM ANALYSIS

Analysis is nothing but finding solution to various problems. System analysis is defined as, process in which we get the information about the existing problems, requirements and to solve various problems related to system. Study of feasibility plays vital role in system analysis, which helps in providing goals for development and design.

4.1 Feasibility Analysis

Considering initial results, a deeper survey or feasibility study is done on existing results. "FEASIBILITY ANALYSIS" is nothing but meeting the requirement of system, to achievedesired goal, working on proposed system and getting detailed information on resources.

There are 8 steps to be considered for Feasibility Analysis:

- Develop a team for a project and assign leader for that team.
- Identify the strength of proposed work.
- Characterize and recognize qualities of proposed work.
- Identify cost of proposed work and show its performance.
- Show system performance.
- Choose best framework.
- Prepare report and submit to administrative

Three key features involve in "FEASIBILITY ANALYSIS" are:

- Financial Feasibility
- Technical Feasibility
- Public Feasibility

4.1.1 FINANCIAL

Financial feasibility is one of key factor of feasibility analysis, which is carried outto scrutinize economic cost of organization, limited fund is endowed by companies for development of system. Hence, the developed project cost is within budget, because of freely available resources and can be accessed by open-source.

4.1.2 TECHNICAL

Technical feasibility is entailed to scrutinize technical performance of system. System which is building should have less demand. Hence, this gives number high of demands asked by the clients. The proposed system is built in such way that there is no harm for theuser as well the MR-SGSO executes very well. The project draws all the technical requirements from open source and resources are used which have higher execution time.

Summary

The objective of this chapter is to know the proposed system is feasible or not. The chaptermainly describe about the various keys of feasibility analysis i.e. financial, technical and public.

CHAPTER 5

SYSTEM DESIGN

Design is an innovative procedure; a successful system is developed through great designs. System design is a process of giving detailed information about the proposed work in a physical format. Different designs are built for development ofsystem, which describes about features, components which are included and how client interact with system.

5.1 System Architecture

System architecture describes about each component of system:

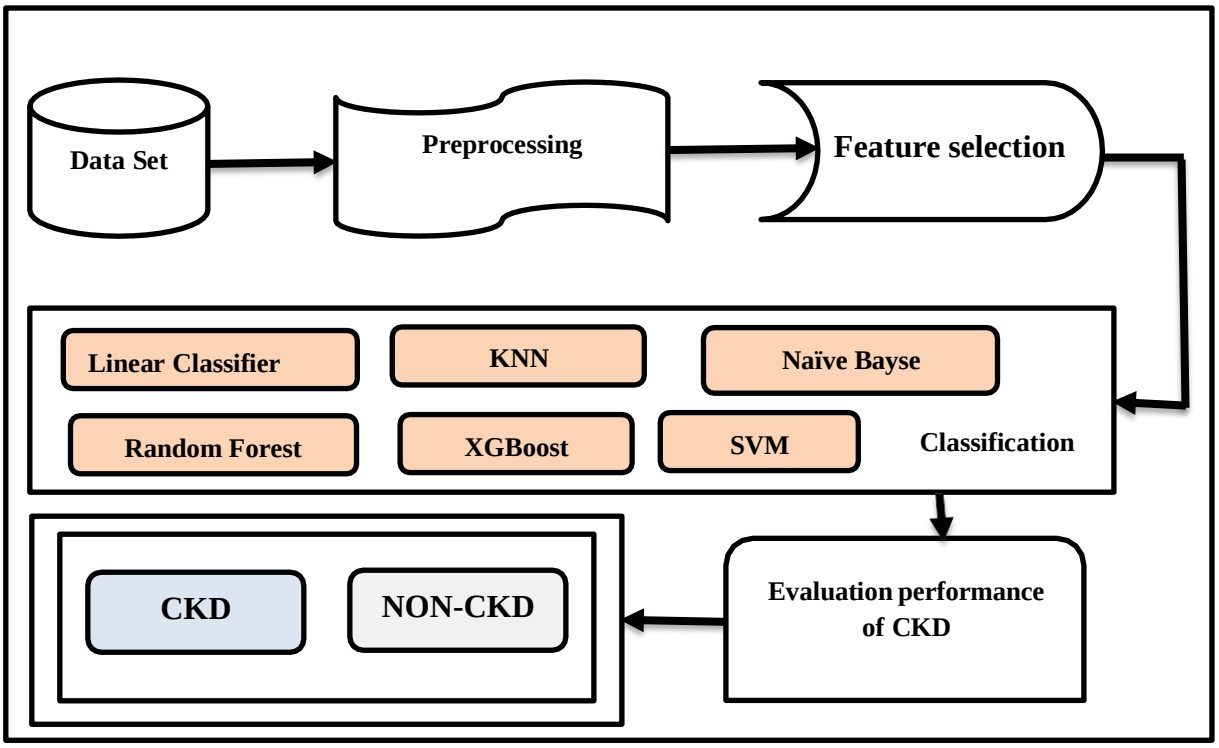

Fig 5.1: System Architecture

5.2 Use case Diagram

A **use case diagram** at its simplest is a representation of a user's interaction with the systemthat shows the relationship between the user and the different use cases in which the user is involved. A use case diagram can identify the different types of users of a system and thedifferent use cases and will often be accompanied by other types of diagrams as well. Theuse cases are represented by either circles or ellipse.

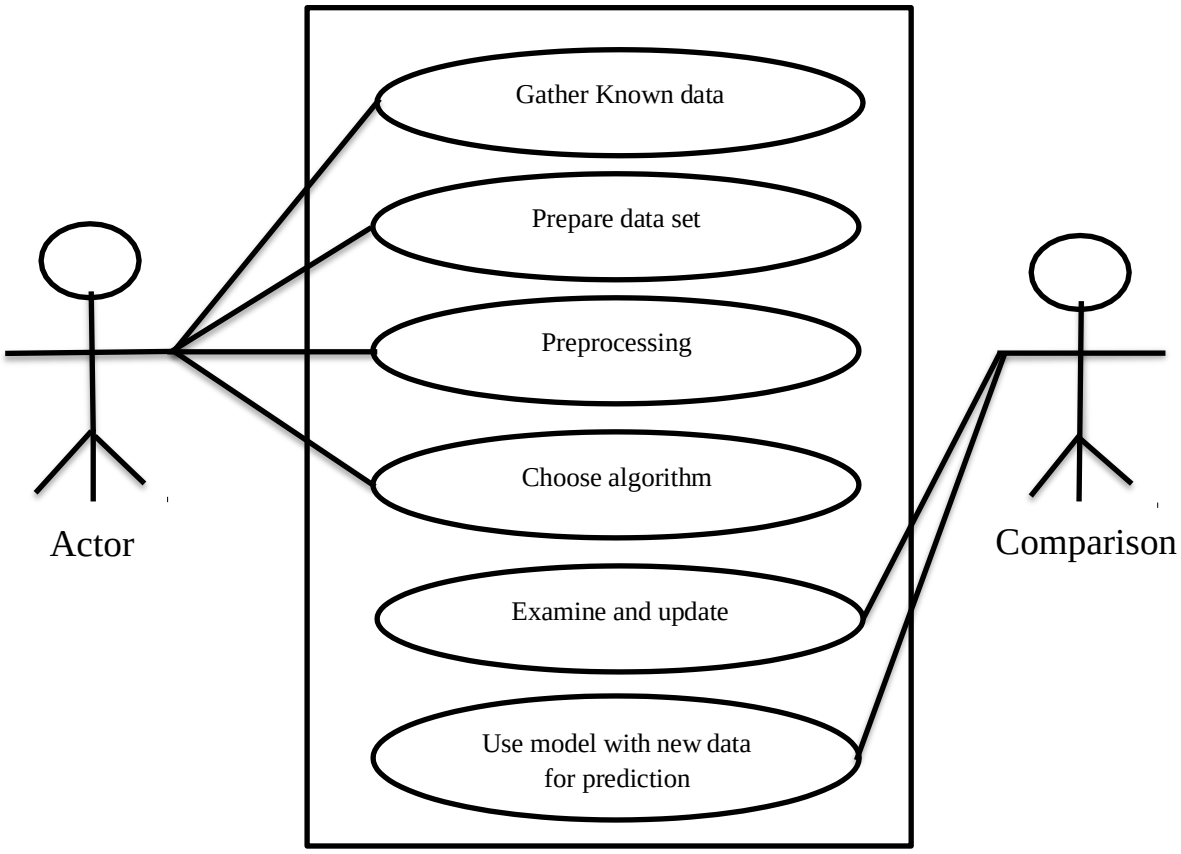

Fig 5.2: Use-case Diagram

5.3 Class Diagram

Class diagrams are generally used for conceptual modeling of static view of a software application, and for modeling translating models into programming code in a detailed manner. At time of developing or construction software systems, a class diagram is widelyused. They are also used for data modeling. It is used to show classes, relationships amongthem, interface, association, etc. Class in a class diagram simply is a blueprint of an object.It simply describes and explains different type of objects in system, and different types of relationships that exist between them.

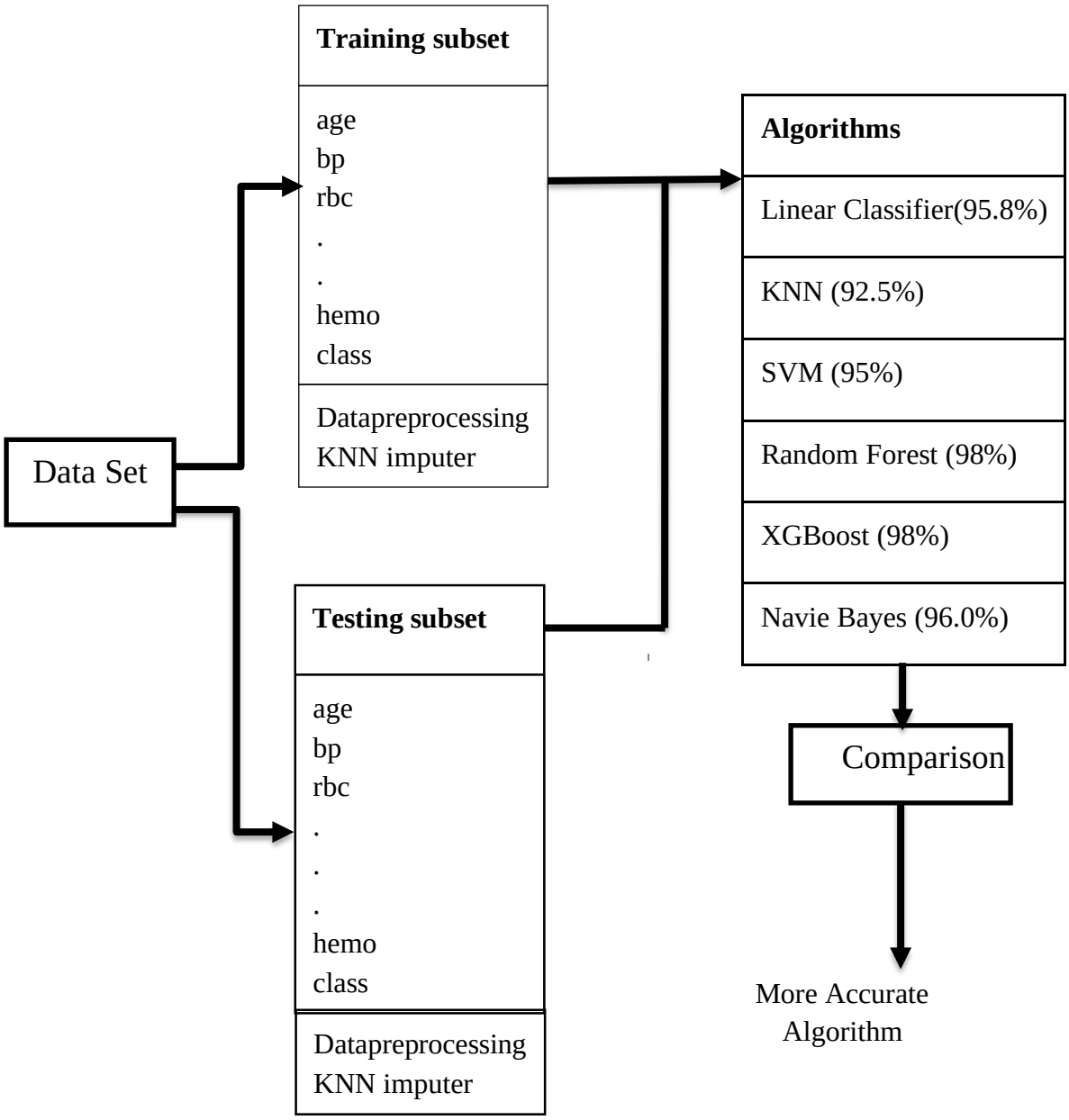

Fig 5.3: Class diagram

Summary

This chapter mainly concentrates on input design, output design, systemarchitecture and various designs which is involved in proposed system.

CHAPTER 6

IMPLEMENTATION

It is main part of project where project is implemented. Implementation must be clearly defined, planned carefully and systematically otherwise it causes confusion and leads to generation of problems. Following tasks involved in stage of implementation:

- Planning carefully
- System Investigation.
- Investigating different types of systems and constraints.
- Selecting the most accurate and correct language for developing the application.
- Understanding and correctly evaluating several changeover methods.
- Rightful decision is made about the selection of the platform.

6.1 Language used for implementation

Implementation is an important step where execution of project is done. Hence language used for implementation should be adaptable, error free, user friendly and etc. Most of thetime because of some mistake project get ruined or spoiled due to inappropriate programming language.

Python is widely considered as the preferred language for ML (Machine Learning). Few simple reasons are:

It has an excellent collection of in-built libraries: Python claims a huge numberof in-built libraries for data mining, data manipulation, and machine learning. For example, NumPy, this is used for scientific calculation. Scikit-learn have tools for data mining and analysis which optimizes the Python's brilliant machine learning usability. Panda is another package which provides developers with high- performance structures and data analysis tools and helps them to reduce the project implementation time. Similarly, SciPy is used for advanced computation, and Pybrain is used for machine learning.

Moderate learning Curve: Python Programming language is very accessible and easy to learn and use. It focuses on code readability. It is versatile and well- structured language.

It is a general-purpose programming language: Well, Python is a good choiceif the project requirements are more than just information such as developing a functional website.

Easy to integrate: Python programming language incorporates better than "R" in business environments. It is easy to integrate Python with other lower-level languages such as C, C++, or Java. Similarly, the Python based-stack is easy to incorporate with data scientist's work, which allows it to bring efficiently into production.

High productivity: Syntax in Python is extremely readable and easy to understand similar to other programming languages, whereas R programminglanguage has different syntax. The readability syntax in Python programming language confirms high productivity of development teams.

Easy to create prototypes: It is already stated that Python is easy to learn and fast to develop. It requires less coding, which means that you can create prototypesand test your concepts quickly and easily in Python as compared to other programming languages. Developing prototypes in Python not only saves developers time, but it also decreases your company's costs.

It is free and open source: Python programming language is available freely; this allows you to download Python for free, i.e., you can download its source code,make the modifications in it and then distribute it. The download comes with an extensive collection of libraries that supports you to carry out your tasks.

Object-oriented paradigm: Python programming language supports for both the Object-oriented and procedural programming models. Classes and objects in Object-oriented programming help us to model the real world while functions in procedural programming enable us to reuse the code. A class in Object-oriented programming encapsulates data and function together.

Portability: Python programming language is portable, i.e. code written in Python can be run on another platform; this is called Write Once Run Anywhere, i.e., WORA, but this is not possible with other languages like C++. When you haveto run C++ code on other platforms, it requires making some modifications in the code to run. It is not advisable to use any system-dependent features.

6.2 Implementation platform

Platform plays an important role for development of software. A platform is a placewhere software is launched. Ubuntu platform is used for implementation of project. Thereare various analyses for choosing Ubuntu platform. It is an open-source operating system.It is used for many different purposes such as scrutiny, embedded system, security and etc.it also contains remote connection and option to restore, so that file can be restored if deleted. Ubuntu software is used for scalable processing, multi-tasking, Encrypted File System (EFS) and smart card support.

6.2.1 Linear Classifier

A linear classifier is a model that makes a decision to categories a set of data points to a discrete class based on a linear combination of its explanatory variables. As an example, combining details about a dog such as weight, height, color and other features would be used by a model to decide its species

A linear classifier will draw a straight line (e.g., y=mx+c) to divide sets of points into 2 classes.

Formula:

$$Y = \beta_0 + \sum_{i=0}^{n} \beta_i . X_i$$

6.2.2 K-Nearest Neighbor Algorithm

It is a supervised ML algorithm used to solve both regression and classification problems.itaims to locate all of the closest neighbors around a data point in order to figure out in whichclass it belongs.

KNN is one of the simplest machine learning algorithms which is easy to implement.

It is a supervised machine learning algorithm that solves both classification and regression problems. This algorithm stores all the available data and classifies new data based on their similarity.

First select the number of neighbors. Then calculate Euclidean distance between the data points.

Euclidean distance is the distance between two points. Formula for Euclian distanceis given by:

$$\sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$$

6.2.3 Support Vector Machine

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for classification problems in Machine Learning.

The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.

SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called support vectors, and hence algorithm is termed a Support Vector Machine. Consider the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane.

Formula used for non-linear data:

$$\mathbf{Z = X^2 + Y^2}$$

6.2.4 Random Forest

Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap and Aggregation, commonly known as **bagging**. The basic idea behind this isto combine multiple decision trees in determining the final output rather than relying on.

industrial decision trees. Random Forest has multiple decision trees as base learning models. We randomly perform row sampling and feature sampling from the dataset forming sample datasets for every model. This part is called Bootstrap.

The random forest algorithm works by completing the following steps:

Step 1: The algorithm selects random samples from the dataset provided.

Step 2: The algorithm will create a decision tree for each sample selected. Then it willget a prediction result from each decision tree created.

Step 3: Voting will then be performed for every predicted result. For a classificationproblem, it will use mode, and for a regression problem, it will use mean.

Step 4: And finally, the algorithm will select the most voted prediction result as the finalprediction.

6.2.5 Naïve Bayes Classifier

Naive Bayes classifiers are a collection of classification algorithms based on **Bayes' Theorem**. It is not a single algorithm but a family of algorithms where all of them sharea common principle, i.e., every pair of features being classified is independent of each other.

Bayes' Theorem finds the probability of an event occurring given the probability of another event that has already occurred. Bayes' theorem is stated mathematically as the following equation:

P (yi | X1, X2, ..., Xn) = P (X1 | yi)P (X2 | yi)...P (Xn | yi)P (yi)/P (X1, X2, ..., Xn)

where A and B are events and P (B) ≠ 0.

- Basically, we are trying to find probability of event A, given the event B is true. EventB is also termed as **evidence**.
- P (A) is the **priori** of A (the prior probability, i.e., Probability of event before evidenceis seen). The evidence is an attribute value of an unknown instance (here, it is event B).
- P (A|B) is a posteriori probability of B, i.e., probability of event after evidence is seen.Now,

with regards to our dataset, we can apply Bayes' theorem in following way:

$$P(y|X) = \frac{P(X|y)\,P(y)}{}$$

where, y is class variable and X is a dependent feature vector (of size *n*) where:

$X = (x_1, x_2, x_3 \ldots\ldots\ldots\ldots x_n).$

6.2.6 XGBoost

XGBoost is an implementation of Gradient Boosted decision trees. XGBoost models majorly dominate in many Kaggle Competitions.

In this algorithm, decision trees are created in sequential form. Weights play an importantrole in XGBoost. Weights are assigned to all the independent variables which are thenfed into the decision tree which predicts results. The weight of variables predicted wrong by the tree is increased and these variables are then fed to the second decision tree. These individual classifiers/predictors then ensemble to give a strong and more precise model.It can work on regression, classification, ranking, and user-defined prediction problems.

Similarity Score = (Sum of residuals)^2 / Number of residuals + lambda

Gain = Left tree (similarity score) + Right (similarity score) - Root (similarity score)Gain - gamma

Output value = Sum of residuals / Number of residuals + lambda

Best Algorithm is chosen based on the following

Precision: Tell us how many of the correctly predicted cases actually turned out to be positive.

$$\frac{TP}{TP+FP}$$

Recall: Tell us how many of the actual positive cases we were able to predict correctly with our model.

$$\frac{TP}{TP+FN}$$

F1-score: harmonic mean of precision & Recall, so it gives a combined idea about these 2metrics. It is maximum when precision is equal to Recall.

Summary

The chapter describes about implementation process of proposed system. It describes aboutwhich platform, language, algorithms and processing techniques used for executing the project.

CHAPTER 7

TESTING

Testing is an important part of project where testing of each module is done. Testing guarantees that proposed system is well organized analyzed to meet require project goal. Testing is last stage of project which guarantees the system is error free and ready to give desired output. The goals of testing are given by:

- Give operational quality to system
- Search and remove errors.
- Best quality project is produced.
- To approve the product as a solution for the first issue.

The following are types of testing performed in proposed system.

7.1 Unit Testing

Each module of project is tested individually. Verification is done on each module. Module of project is tested individually. Testing is done in programming style. The unit testing for the proposed system is performed on initialization, Data collection, Genre classification, visualizing the output.

7.2 Integration

After unit testing is performed, integration testing takes place in each module of project. Integration is done on various classes of system. It is done on front and back end also.

- **Function into classes Integration**

  Initially during code phase various functions is developed for development of system. Each function of system is tested and coded individually. As all the functions are verified, they are mixed into their particular classes.

- **Distinct classes Integration**

  Based on functionality, testing of distinct classes is done independently. Verification of each class is done which gives good result and hence integration is performed again on different classes.

7.3 Integration Testing

Developing a programming framework is a sophisticated technique which is used by Integration testing. It solves various issues on dual verification problem and construct program which solves all related problems. Main objective of integration testing is to construct a program structure based on unit testing modules.

As modules of software are divided, testing is performed on each module. Later this separated module is tested as whole set. Here, to rectify errors is difficult as it has differentisolated errors.

7.3.1 Top-down Integration

The up down integration deals with development of program framework in incremental way. Modules of each program are coordinated in descending order and it startswith primary module. This method is an incremental approach to the construction of program structure. Modules are integrated by moving downward, beginning with the main program module. Modules that subordinate to the main program module are incorporatedinto the structure in either a depth first or breadth first manner.

7.3.2. Bottom-up Integration

The below table integrated testing table is divided into integrated classes, functionsof each classes, how test is performed and result generated. This method begins the construction and testing with the modules at the lowest level in the program structure. Sincethe modules are integrated from bottom to up, processing required for modules subordinateto a given level is always available. Therefore, in this case the need for stubs is eliminated.It is important to check whether the testing is error free or not for different classes.

7.4 Validation Testing

As completion of testing combination is done, writing computer program is put together inon package. Testing approval is described from various perspectives. Hence testing affirmsitems limits which are sensibly expecte

7.5 User Acceptance Testing:

Key factor for success of any system is user acceptance. User acceptance testing is performed on users, show which will be successful based on user motivation and knowledge. At time of developing and making required changes, system under consideration along with prospective system, users undergo constant testing. The changes are made regarding to 3 points

- Input design
- Output design

7.6 Artificial Data

Artificial data is created for testing purpose. All combinations of formats and values are tested by generating the artificial test data. Data generating utility program in system information departments helps to prepare artificial data as quickly as possible. It is made possible to test all control paths, login through program.

7.7 Assurance of Quality

Quality assurance is testing and analysis of administration element. The main objective is to give knowledge about item quality to administration. Assurance of quality involves:

7.7.1 Quality Factors

The main goal of confirmation value is track product quality and observes procedures to enhance programming.

Quality factors are described in following two categories:

- Directly measured factors.
- Indirectly measured factors.

Quality factors mainly focus on following three things:

- Operational attributes
- Experiences capacity changes
- Versatility.
- Effectiveness
- Time duration.

7.7.2 General Risks

Risk is nothing but which gives negative results at undesirable incident. The following threethings are considered to recognize risk in other projects:

- Damage occurs during an occasion.
- Probability of occurring an occasion.
- The result that modified at certain level.

7.7.3 Security Technologies & Policies

The seven major activities the software quality is comprised of which are follows:

- Conduct of formal specialized audits
- Measurement
- Testing Software
- Application for specialized techniques.
- Control of progress
- Record keeping and announcing
- Enforcement of measures

Summary

The chapter describes about various testing techniques such as unit testing, integration testing, validation testing, output testing and etc.

CHAPTER 8

RESULTS

8.1 Dataset

The Dataset which has been use in this project is Chronic Kidney Disease Dataset. ThisDataset contains 400 records of different parameter from patients and non-patients.

8.2 Results

8.2.1 CONFUSION MATRIX OF LINEAR CLASSIFIER

True Negative	**TN**	00	69
True Positive	TP	11	46
False Negative	FN	10	04
False Positive	FP	01	01

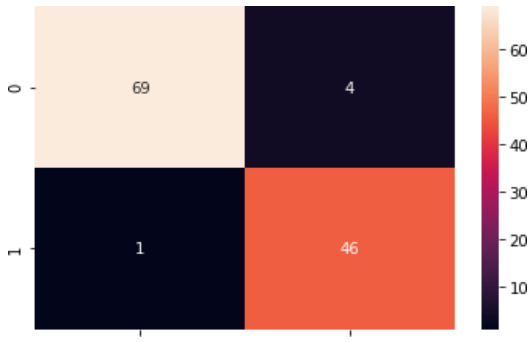

Fig 8.2.1: Confusion matrix of Linear Classifier

Calculation:

$$\frac{TP+TN}{TP+TN+FN+FP} = \frac{46+69}{46+69+04+01} = \frac{115}{120} = 0.9583$$

CLASSIFICATION REPORT OF LINEAR CLASSIFIER

	Precision	**Recall**	**F1-score**	**Support**
0	**0.99**	**0.95**	**0.97**	**73**
1	**0.92**	**0.98**	**0.95**	**47**
Accuracy			**0.96**	**120**
Macro average	**0.95**	**0.96**	**0.96**	**120**
Weighted average	**0.96**	**0.96**	**0.96**	**120**

Table 8.2.1 Classification report of Linear classifier

8.2.2 CONFUSION MATRIX OF KNN

True Negative	**TN**	00	58
True Positive	**TP**	11	53
False Negative	**FN**	10	05
False Positive	**FP**	01	04

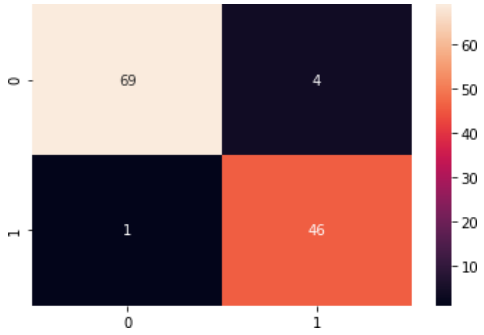

Fig 8.2.2: Confusion matrix of KNN

Calculation:

$$\frac{TP+TN}{TP+TN+FN+FP} = \frac{53+58}{53+58+05+04} = \frac{111}{120} = 0.925$$

CLASSIFICATION REPORT OF KNN

	Precision	**Recall**	**F1-score**	**Support**
0	**0.94**	**0.92**	**0.93**	**63**
1	**0.91**	**0.93**	**0.92**	**57**
Accuracy			**0.93**	**120**
Macro average	**0.92**	**0.93**	**0.92**	**120**
Weighted average	**0.93**	**0.93**	**0.93**	**120**

Table 8.2.2 Classification report of KNN

8.2.3 CONFUSION MATRIX OF SVM

True Negative	**TN**	00	65
True Positive	**TP**	11	51
False Negative	**FN**	10	03
False Positive	**FP**	01	01

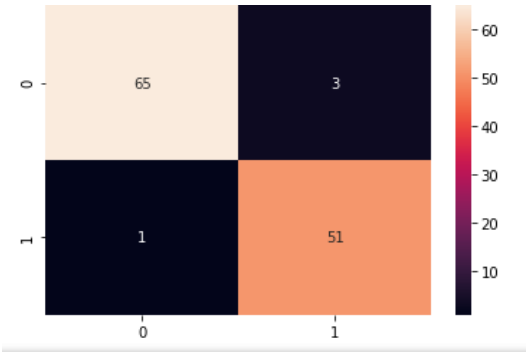

Fig 8.2.3: Confusion matrix of SVM

Calculation:

$$\frac{TP+TN}{TP+TN+FN+FP} = \frac{51+65}{51+65+03+01} = \frac{116}{120} = 0.966$$

CLASSIFICATION REPORT OF SVM

	Precision	**Recall**	**F1-score**	**Support**
0	**0.98**	**0.96**	**0.97**	**68**
1	**0.94**	**0.98**	**0.96**	**52**
Accuracy			**0.97**	**120**
Macro average	**0.96**	**0.97**	**0.97**	**120**
Weighted average	**0.97**	**0.97**	**0.97**	**20**

Table 8.2.3 Classification report of SVM

8.2.4 CONFUSION MATRIX OF Random Forest

True Negative	**TN**	00	66
True Positive	**TP**	11	32
False Negative	**FN**	10	00
False Positive	**FP**	01	02

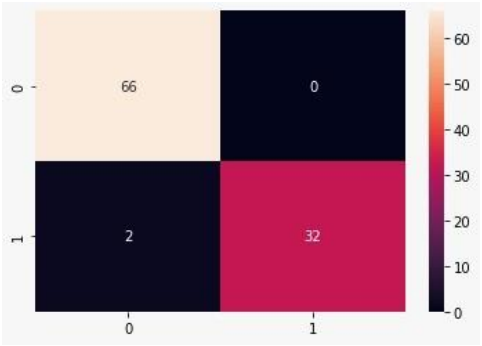

Fig 8.2.4: Confusion matrix of Random Forest

Calculation:

$$\frac{TP + TN}{TP+TN+FN+FP} = \frac{32 + 66}{32+66+00+02} = \frac{98}{100} = 0.98$$

CLASSIFICATION REPORT OF RANDOM FOREST

	Precision	**Recall**	**F1-score**	**Support**
0	**0.94**	**1.00**	**0.97**	**66**
1	**1.00**	**0.88**	**0.94**	**34**
Accuracy			**0.96**	**100**
Macro average	**0.97**	**0.94**	**0.95**	**100**
Weighted average	**0.96**	**0.96**	**0.96**	**100**

Table 8.2.4 Classification report of Random Forest

8.2.5 CONFUSION MATRIX OF XGBOOST

True Negative	**TN**	00	66
True Positive	**TP**	11	32
False Negative	**FN**	10	00
False Positive	**FP**	01	02

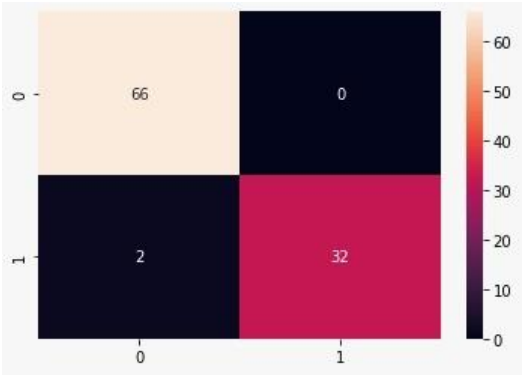

Fig 8.2.5: confusion matrix of XGBoost

Calculation:

$$\frac{TP + TN}{TP+TN+FN+FP} = \frac{32 + 66}{32+66+00+02} = \frac{98}{100} = 0.98$$

CLASSIFICATION REPORT OF XGBoost

	Precision	**Recall**	**F1-score**	**Support**
0	**0.97**	**1.00**	**0.99**	**66**
1	**1.00**	**0.94**	**0.97**	**34**
Accuracy			**0.98**	**100**
Macro average	**0.99**	**0.97**	**0.98**	**100**
Weighted average	**0.98**	**0.98**	**0.98**	**100**

Table 8.2.5 Classification report of XGBoost

8.2.6 CONFUSION MATRIX OF NAÏVE BAYES

True Negative	**TN**	00	62
True Positive	**TP**	11	34
False Negative	**FN**	10	4
False Positive	**FP**	01	0

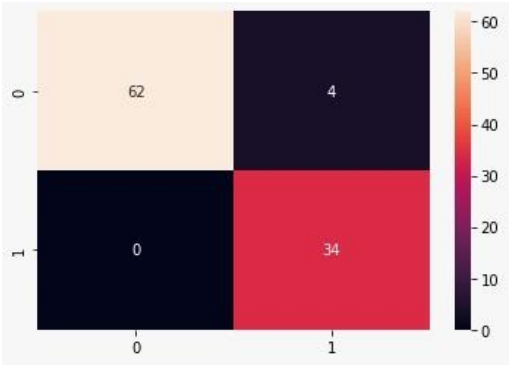

Fig 8.2.6: Confusion matrix of Naïve Bayes

Calculation:

$$\frac{TP + TN}{TP+TN+FN+FP} = \frac{34+ 62}{34+62+04+00} = \frac{96}{100} = 0.96$$

CLASSIFICATION REPORT OF NAÏVE BAYES

	Precision	**Recall**	**F1-score**	**Support**
0	**1.00**	**0.94**	**0.97**	**66**
1	**0.89**	**1.00**	**0.94**	**34**
Accuracy			**0.96**	**100**
Macro average	**0.95**	**0.97**	**0.96**	**100**
Weighted average	**0.96**	**0.96**	**0.96**	**100**

Table 8.2.6 Classification report of Naïve Bayes

8.3 COMPARISION TABLE

	Accuracy (%)	precision (%)	recall (%)	F1-score (%)
Linear Classifier	96.0	100	88.2	93.7
KNN	96.0	96.8	91.1	93.9
SVM	95.0	93.9	91.1	92.5
XGBoost	98.0	100	94.11	92.5
Random Forest	98.0	100	94.11	96.9
Naïve Bayes	96.0	89.4	100	94.4

Table 8.3.1: Comparison of Algorithms

8.4 COMPARISION GRAPHS

Accuracy

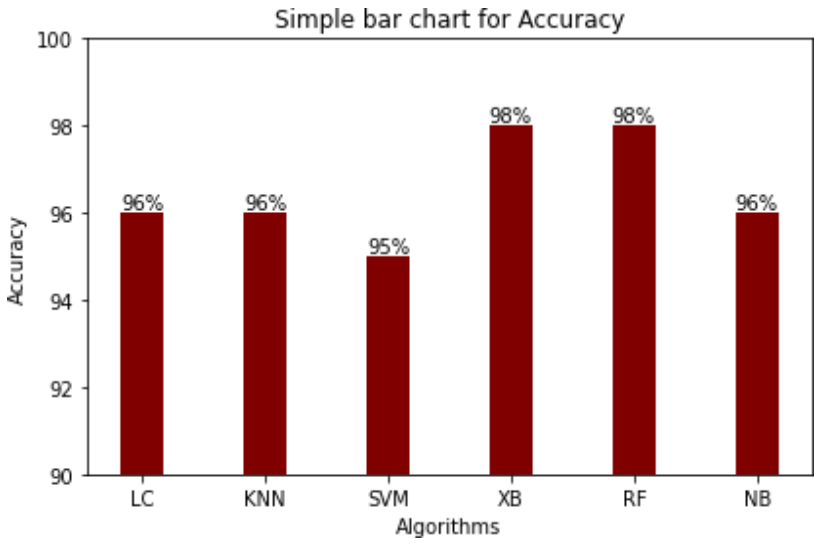

Fig 8.4.1: Accuracy Bar Chart

Precission

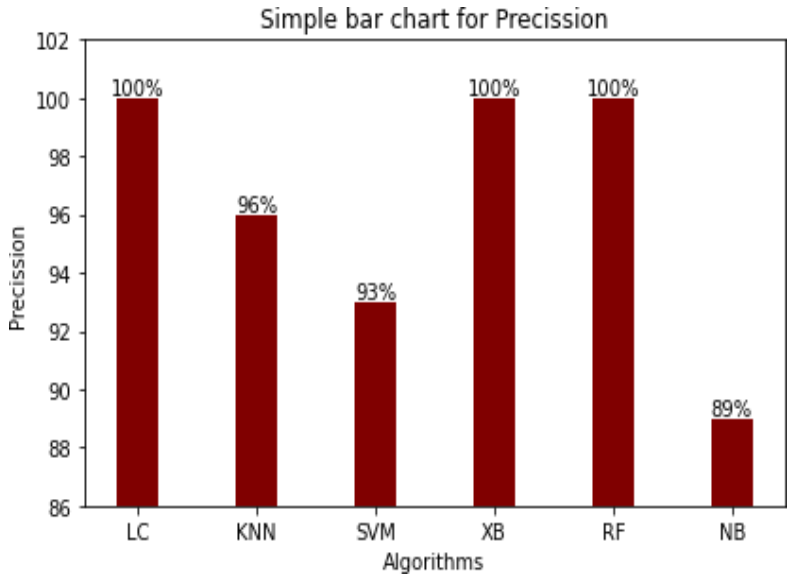

Fig 8.4.2: Precission Bar chart

Recall

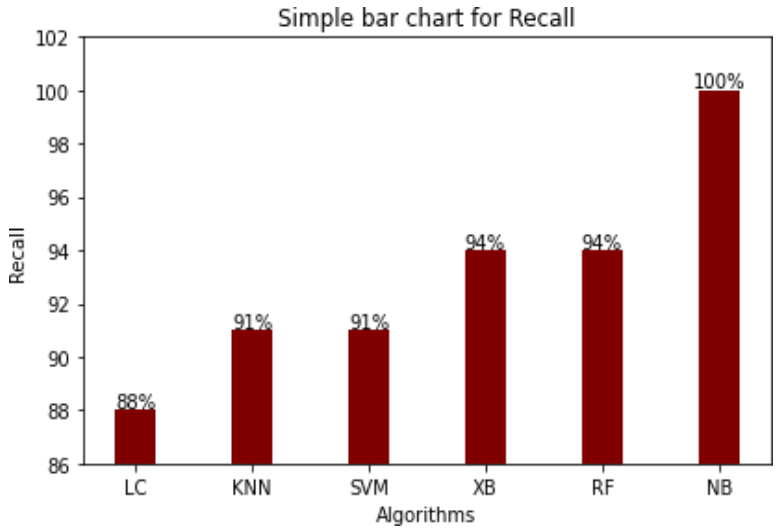

Fig 8.4.3: Recall Bar chart

F1-Score

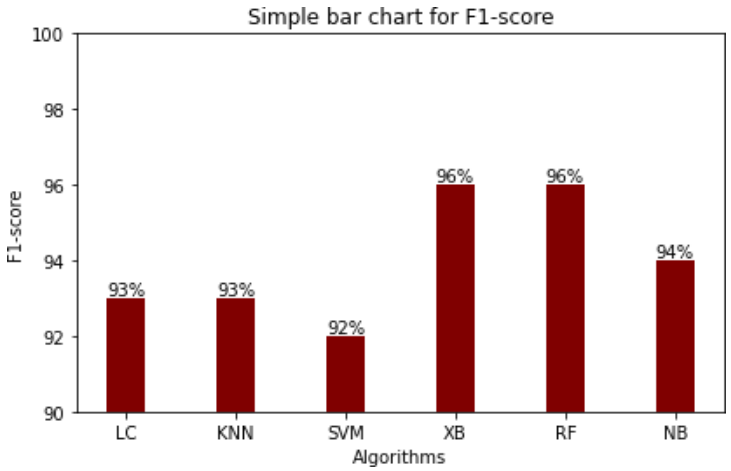

Fig 8.4.4: F1-score Bar chart

Summary

The chapter describes about various screenshot of results, step by step process toexecute proposed project and advantages of proposed system

CONCLUSION

One of the disorders that affects a lot of individuals is chronic kidney disease (CKD). The disease is frequently only discovered once it has progressed to an advanced state because the symptoms of CKD are not noticeable in the early stages. This could result in kidney failure and death. Machine learning classifiers offer an effective means of early disease prediction. The Linear Classifier, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), XGBoost, Naive Bayes, and Random Forest machine learning techniques were utilised. Several measures were used to assess how well these classifiers performed. The XGBoost fared better with 98 percent accuracy. However, as the dataset was a little unbalanced, accuracy cannot be the only factor taken into account when evaluating the parameter. Based on Precision, XGBoost and Random Forest both displayed 100%, whereas Linear Classifier also displayed 100%. Based on recall, XGBoost demonstrated 94%, Random Forest demonstrated 94%, and Naive Bayes demonstrated 100%. Based on the F1 score, XGBoost and Random Forest both displayed 96 percent. After examining every metric, we can say that accuracy is a good indicator if the datasets are fairly balanced and all types of outputs are valued equally. However, F1-score combines precision and recall and can be used in cases where the datasets are unbalanced because it requires both precision and recall to have a reasonable value. The formula will weigh the metric value down if the precision or recall of the positive class is low, even if there are few positive cases compared to negative cases. So, we are concluding our accurate algorithm by comparing with the metric F1-score. Based on F1-score XGBoost and Random Forest showed 96%. By carefully analyzing about the bagging and boosting algorithms, boosting algorithms can be more precise than the bagging algorithms. Based on the evaluation XG Boost was the best classifier when compared with all the other algorithms.

Therefore, it is important to investigate how the KNN imputer-based method may be used in the future to handle missing values in data sets relating to various diseases.

REFERENCES

[1] Akanksha Soni and Dr. Avinash Rai "Kidney stone recognition and extraction using directional Emboss & SVM from Computes Tomography Images" in 2020 Third International Conference on Multimedia Processing, Communication & Information Technology – MPCIT 2020 JNNCE, Shivamogga, Karnataka, India DOI:10.1109/MPCIT51588.2020.9350388

[2] M.Akshaya, R.Nithushaa, N.Sri Madhava Raja and S.Padmapriya "Kidney Stone Detection Using Neural Networks" in 2020 International conference on system, computation, automation, and networking (ICSCAN). DOI:10.1109/ICSCAN49426.2020.9262335

[3] Gabriel R. Vasquez- Morales, Sergio M. Martinez-Monterrubio, Pablo Moreno- Ger, and Juan A. Recio-Garcia, ''Explainable Prediction of Chronic Renal Disease in the Colombian Population Using Neural Networks and Case-Based Reasoning,'' in 2019 IEEEAccess DOI: 10.1109/ACCESS.2019.2945430

[4] L. Antony et al., "A Comprehensive Unsupervised Framework for Chronic Kidney Disease Prediction," in IEEE Access, vol. 9, pp. 126481-126501, 2021, DOI: 10.1109/ACCESS.2021.3109168.

[5] Pankaj chittora, Sandeepchaurasia,Prasunchakrabatri, Guaravkumawat "Chronic Kidney disease prediction using Machine Learning," in 2021 IEEE Access 2021 DOI: 10.1109/ACCESS.2021.3053763

[6] Nilar Thein, Hanung Adi Nugroho, Teguh Bharata Adji and Kazuhiko Hamamoto"An image pre-processing method for kidney stone segmentation in CT scan images" in 2018 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM) DOI: 10.1109/CENIM.2018.8710933

[7] Saranya.N, Sakthi Samyuktha M, Sharon Isaac, Subhanki B "Diagnosing Chronic Kidney Disease using KNN Algorithm" in 2021 DOI: 10.1109/ICACCS51430.2021.9442039

[8] Sunghwan Kim "An Empirical Evaluation of Machine Learning Techniquesfor Chronic Kidney Disease Prophecy" in 2020 IEEE Access 2020 DOI: 10.1109/ACCESS.2020.2981689

[9] N. S. Altman, "An introduction to kernel and nearest-neighbour non parametric regression," The American Statistician, vol. 46, no. 3, pp. 175–185, 1992. DOI: 10.1109/ICCCNT45670.2019.8944799

[10] U. Ekanayake and D. Herath, "Chronic Kidney Disease Prediction Using Machine

Learning Methods," 2020 Moratuwa Engineering Research Conference (MERCon), 2020, pp. 260-265, doi: 10.1109/MERCon50084.2020.9185249.

[11] Bhavya Gudeti, Shashvi Mishra, Shaveta Malik, Terrance Frederick Fernandez, Amit Kumar Tyagi, Shabnam Kumari "A Novel Approach to Predict Chronic Kidney Disease using Machine Learning Algorithms"

[12] R. Praveen Kumar, Sarath Erive, Ganta Jayasri, Annapalli Srujana, VallalaNiharika, "Predicting Chronic Kidney Disease by Applying Feature Engineering & Performance Analysis of Machine Learning Classifiers" in 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA) DOI: 10.1109/ICECA52323.2021.9675935

[13] Yedilkhan Amirgaliyev, Shahriar Shamiluulu, Azamat Serek, "Analysis of Chronic Kidney Disease Dataset by Applying Machine Learning Methods" in 2018 IEEE 12th International Conference on Application of Information and Communication Technologies (AICT), DOI: 10.1109/ICAICT.2018.8747140

[14] Shubham Vashisth, Ishika Dhall, Shipra Saraswat, "Chronic Kidney Disease (CKD) Diagnosis using Multi-Layer Perceptron Classifier", in 2020 10th InternationalConference on Cloud Computing, Data Science & Engineering (Confluence), DOI: 10.1109/Confluence47617.2020.9058178

[15] A. Vijayalakshmi, V. Sumalatha, "Survey on Diagnosis of Chronic KidneyDisease UsingMachine Learning Algorithms" in 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), DOI: 10.1109/ICISS49785.2020.9315880

[16] Eshtiak Ahmed, Asif Karim, Sami Azam, "Chronic Renal Disease Prediction using Clinical Data and Different Machine Learning Techniques" in 2021 2nd International Informatics and Software Engineering Conference (IISEC), DOI: 10.1109/IISEC54230.2021.9672365

[17] Faisal Arafat, Thaharim Khan, Atanu Das Bapon, Md. Ibrahim Khan, Sheak Rashed Haider Noori, "A Deep Learning Approach to Predict Chronic Kidney Disease in Human" in 2021 IEEE 12th Annual Information Technology, Electronics and MobileCommunication Conference (IEMCON). DOI: 10.1109/IEMCON53756.2021.9623101

[18] Md. Rashed-Al-Mahfuz, Abedul Haque, Akm Azad, Salem A. Alyami, Julian M. W. Quinn, Mohammad Ali Moni, "Clinically Applicable Machine Learning Approaches to Identify Attributes of Chronic Kidney Disease (CKD) for Use in Low-Cost Diagnostic Screening" in 2021 IEEE Journal of Translational Engineering in Health and Medicine (Volume: 9), DOI: 10.1109/JTEHM.2021.3073629

[19] Zobair Hasan, Rafiur Rahman Khan, Wazed Rifat, Dipok Sarkar Dipu, Muhammad Nazrul Islam, Iqbal H. Sarker, "Development of a Predictive Analytic System for Chronic Kidney Disease using Ensemble-based Machine Learning" in 2021 62nd International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS), DOI: 10.1109/ITMS52826.2021.9615273

[20] M.P.N.M. Wickramasinghe, D.M. Perera, K.A.D.C.P. Kahandawaarachchi, "Dietary prediction for patients with Chronic Kidney Disease (CKD) by considering blood potassium level using machine learning algorithms" in 2017 IEEE Life Sciences Conference (LSC), DOI: 10.1109/LSC.2017.8268202

[21] Fadil Iqbal, Aruna S. Pallewatte, Janaka P. Wansapura, "Texture analysis ofultrasound images of chronic kidney disease" in 2017 Seventeenth International Conference on Advances in ICT for Emerging Regions (ICTer), DOI: 10.1109/ICTER.2017.8257787

[22] Uma N Dulhare, Mohammad Ayesha, "Extraction of action rules for chronic kidney disease using Naïve bayes classifier" in 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), DOI: 10.1109/ICCIC.2016.7919649

[23] Rahul Gupta, Nidhi Koli, Niharika Mahor, N Tejashri, "Performance Analysis of Machine Learning Classifier for Predicting Chronic Kidney Disease" in 2020 International Conference for Emerging Technology (INCET), DOI: 10.1109/INCET49848.2020.9154147

[24] Shubham Vashisth, Ishika Dhall, Shipra Saraswat, "Chronic Kidney Disease (CKD) Diagnosis using Multi-Layer Perceptron Classifier" in 2020 10th International

Conference on Cloud Computing, Data Science & Engineering (Confluence), DOI: 10.1109/Confluence47617.2020.9058178

[25] Erni Yudaningtyas, Djoko H. Santjojo, Waru Djuriatno, Indrazno Siradjuddin, Muhammad Rony Hidayatullah, "Identification of pulse frequency spectrum of chronickidney disease patients measured at TCM points using FFT processing" in 2017 15th International Conference on Quality in Research (QiR): International Symposium on Electrical and Computer Engineering, DOI: 10.1109/QIR.2017.8168476