

# **DAYANANDA SAGAR UNIVERSITY**

**KUDLU GATE, BANGALORE – 560068**



**Bachelor of Technology  
in  
COMPUTER SCIENCE AND ENGINEERING**

**Major Project Report  
On  
CHRONIC KIDNEY DISEASE PREDICTION USING  
MACHINE LEARNING**

**By**

**Abhigna Srikara G - ENG19CS0008**

**Akhila - ENG19CS0020**

**Anjana KP - ENG19CS0036**

**K Vaishali - ENG19CS0134**

**Under the supervision of**

**Prof. Ranjini K**

**Assistant Professor, Computer Science and Engineering**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING,  
SCHOOL OF ENGINEERING  
DAYANANDA SAGAR UNIVERSITY  
2023**



**DAYANANDA SAGAR UNIVERSITY**

**School of Engineering**  
**Department of Computer Science & Engineering**

Kudlu Gate, Bangalore – 560068  
Karnataka, India

**CERTIFICATE**

This is to certify that the Major project work titled “**CHRONIC KIDNEY DISEASE PREDICTION USING MACHINE LEARNING**” is carried out by **Abhigna Srikara G (ENG19CS0008), Akhila (ENG19CS0020), Anjana KP (ENG19CS0036), K Vaishali (ENG19CS0134)**, bonafide students of Bachelor of Technology in Computer Science and Engineering at the School of Engineering, Dayananda Sagar University, Bangalore in partial fulfillment for the award of degree in Bachelor of Technology in Computer Science and Engineering, during the year **2022-2023**.

**Prof Ranjini K**

Assistant Professor  
Dept. of CS&E,  
School of Engineering  
Dayananda Sagar University

Date:

**Dr. Girisha G S**

Chairman CSE  
School of Engineering  
Dayananda Sagar University

Date:

**Dr. Udaya Kumar  
Reddy K R**

Dean  
School of Engineering  
Dayananda Sagar  
University

Date:

**Name of the Examiner**

- 1.
- 2.

**Signature of Examiner**

## DECLARATION

We, **Abhigna Srikara G (ENG19CS0008), Akhila (ENG19CS0020), Anjana KP (ENG19CS0036), K Vaishali (ENG19CS0134)**, are students of eighth semester B. Tech in **Computer Science and Engineering**, at School of Engineering, **Dayananda Sagar University**, hereby declare that the Major Project Stage-II titled “**CHRONIC KIDNEY DISEASE PREDICTION USING MACHINE LEARNING**” has been carried out by us and submitted in partial fulfillment for the award of degree in **Bachelor of Technology in Computer Science and Engineering** during the academic year **2022-2023**.

Students

Signature

Abhigna Srikara G

ENG19CS0008

Akhila

ENG19CS0020

Anjana KP

ENG19CS0036

K Vaishali

ENG19CS0134

Place: Bangalore

Date:

## ACKNOWLEDGEMENT

*It is a great pleasure for us to acknowledge the assistance and support of many individuals who have been responsible for the successful completion of this project work.*

*First, we take this opportunity to express our sincere gratitude to the School of Engineering & Technology, Dayananda Sagar University for providing us with a great opportunity to pursue our Bachelor's degree in this institution.*

*We would like to thank **Dr. Udaya Kumar Reddy K R, Dean, School of Engineering & Technology, Dayananda Sagar University** for his constant encouragement and expert advice.*

*It is a matter of immense pleasure to express our sincere thanks to **Dr. Girisha G S, Department Chairman, Computer Science and Engineering, Dayananda Sagar University**, for providing right academic guidance that made our task possible.*

*We would like to thank our guide **Prof Ranjini K, Assistant Professor, Dept. of Computer Science and Engineering, Dayananda Sagar University**, for sparing his/her valuable time to extend help in every step of our project work, which paved the way for smooth progress and fruitful culmination of the project.*

*We would like to thank our **Project Coordinator Dr. Meenakshi Malhotra and Dr. Pramod Kumar Naik** as well as all the staff members of Computer Science and Engineering for their support.*

*We are also grateful to our family and friends who provided us with every requirement throughout the course.*

*We would like to thank one and all who directly or indirectly helped us in the Project work.*

# TABLE OF CONTENTS

	Page
LIST OF ABBREVIATIONS.....	vii
LIST OF FIGURES .....	viii
ABSTRACT.....	ix
CHAPTER 1 INTRODUCTION.....	1
1.1 SCOPE .....	4
1.2. SOCIETAL/ENVIRONMENTAL IMPACT .....	4
CHAPTER 2 PROBLEM DEFINITION .....	5
CHAPTER 3 LITERATURE SURVEY .....	7
CHAPTER 4 PROJECT DESCRIPTION.....	12
4.1. PROPOSED DESIGN.....	13
4.1.1. ARCHITECTURE.....	13
CHAPTER 5 REQUIREMENTS .....	14
5.1. FUNCTIONAL REQUIREMENTS .....	15
5.2. NON-FUNCTIONAL REQUIREMENTS .....	15
5.3. HARDWARE AND SOFTWARE REQUIREMENTS .....	15
CHAPTER 6 METHODOLOGY.....	16
6.1. MODULES .....	17
6.1.1. DATA COLLECTION.....	17
6.1.2. DATASET DESCRIPTION .....	18
6.1.3. DATA PREPARATION.....	18
6.1.4. MODEL SELECTION .....	19
6.1.5. ANALYZE AND PREDICTION .....	21
CHAPTER 7 EXPERIMENTATION .....	22
7.1. AGE VS ALL CONTINUOUS COLUMNS.....	23
7.2. HEAT MAP.....	24
7.3. SPLITTING THE DATASET.....	25
7.4. MACHINE LEARNING MODEL.....	25
CHAPTER 8 TESTING AND RESULTS .....	26
8.1. RESULTS .....	27

8.1.1 DATASET UPLOAD .....	27
8.1.2 PREVIEW OF DATASET .....	28
8.2. TEST CASES .....	29
8.2.1. CASE 1: WITH CKD .....	29
8.2.2. CASE 2: WITHOUT CKD .....	30
8.3. ANALYSIS .....	31
8.3.1. BAR GRAPH .....	31
8.3.2. PIE CHART .....	32
CHAPTER 9 CONCLUSION AND FUTURE WORK.....	33
9.1. CONCLUSION .....	34
9.2. SCOPE FOR FUTURE WORK .....	34
CHAPTER 10	
REFERENCES.....	35
APPENDIX .....	37
SAMPLE CODE.....	38
Funding and Published Paper details.....	46

## LIST OF ABBREVIATIONS

CKD	CHRONIC KIDNEY DISEASE
ML	MACHINE LEARNING
GFR	GLOMERULAR FILTRATION RATE
CFS	CORRELATION BASED FEATURE SELECTION
KNN	K NEAREST NEIGHBOURS
LR	LOGISTIC REGRESSION
RF	RANDOM FOREST
SVM	SUPPORT VECTOR MACHINE
DT	DECISION TREE CLASSIFIER

## LIST OF FIGURES

<b>Fig. No</b>	<b>Description of Figures</b>	<b>Page No</b>
1.1	Theoretical flow of CKD	3
4.1.1	Architecture Design	13
6.1.4.1	Logistic Regression	19
6.1.4.2	Random Forest	20
7.1	Age vs continuous column	23
7.2	Finding correlation Between Independent Columns	24
7.3	Splitting the data into Train and Test	25
7.4	Building a Machine Learning Model	25
8.1.1.	Dataset Upload Page	27
8.1.2.1	Preview Of Dataset	28
8.1.2.2	End of Dataset and Train	28
8.2.1	Patient with CKD	29
8.2.2	Patient without CKD(Normal)	30
8.3.1	Bar Graph	31
8.3.2	Pie chart	32



## ABSTRACT

Chronic kidney disease (CKD) is a global health problem with high morbidity and mortality rate, and it induces other diseases. Since there are no obvious symptoms during the early stages of CKD, patients often fail to notice the disease. Early detection of CKD enables patients to receive timely treatment to ameliorate the progression of this disease. Machine learning models can effectively aid clinicians achieve this goal due to their fast and accurate recognition performance. In this study, we propose a machine learning methodology for diagnosing CKD. The CKD data set was obtained from the University of California Irvine (UCI) machine learning repository, which has a large number of missing values. KNN imputation was used to fill in the missing values, which selects several complete samples with the most similar measurements to process the missing data for each incomplete sample. Missing values are usually seen in real-life medical situations because patients may miss some measurements for various reasons. After effectively filling out the incomplete data set, six machine learning algorithms (logistic regression, random forest, support vector machine, k-nearest neighbour, naive Bayes classifier and feed forward neural network) were used to establish models. Among these machine learning models, random forest achieved the best performance with 99.75% diagnosis accuracy. By analyzing the misjudgments generated by the established models, we proposed an integrated model that combines logistic regression and random forest by using perceptron, which could achieve an average accuracy of 99.83% after ten times of simulation. Hence, we speculated that this methodology could be applicable to more complicated clinical data for disease diagnosis.

# **CHAPTER 1**

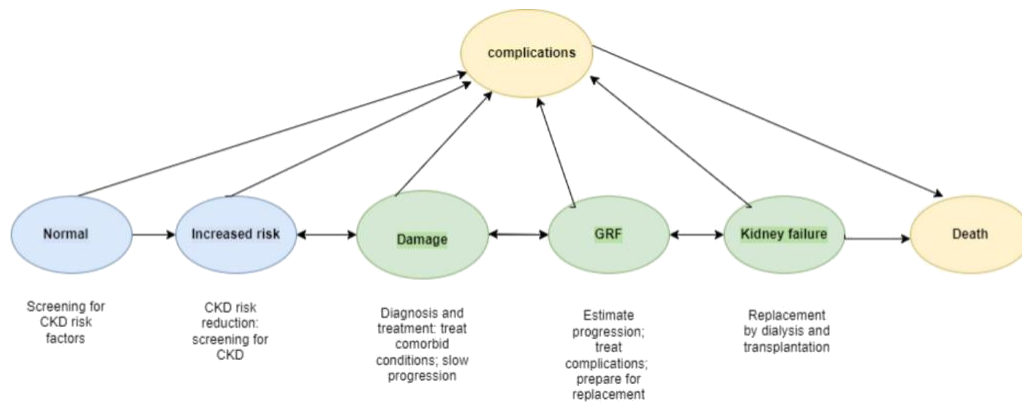
## **INTRODUCTION**

# CHAPTER 1 INTRODUCTION

Chronic kidney disease (CKD) is a global health problem with high morbidity and mortality rate, and it induces other diseases. Since there are no obvious symptoms during the early stages of CKD, patients often fail to notice the disease. Early detection of CKD enables patients to receive timely treatment to ameliorate the progression of this disease. Chronic kidney disease is the loss of kidney function. Oftentimes, the symptoms of the disease are not noticeable and a significant amount of lives are lost annually due to the disease. It is a progressive harmful disease that is propagated in different situations. With the right tools, however, it is possible to make the correct decision in order to do the proper treatment.

Some of the procedures to assess the severity of CKD is by doing tests on both blood and urine as it might be employed for early identification of CKD because it is defined mostly by the presence of excessive albuminuria or eGFR. Kidneys are healthy if eGFR value is more than 90. If eGFR is less than 60, it can be diagnosed as CKD. The patient is diagnosed with the end-stage renal disease if the eGFR result is less than 15. At this point in time the only viable treatments instead of avoiding or postponing kidney failure, there may be additional effective therapies available, such as Kidney donor transplantation and Hemodialysis or Renal replacement therapy. Doctors also use the glomerular filtration rate (GFR) to predict kidney disease. Criteria to define chronic kidney disease are renal impairment with or without decreased GFR for 3 months or (GFR) of 60 mL for 3 months with or without renal impairment Less than 1.73 m<sup>2</sup>. The most reliable indicator of renal function for identifying the various stages of CKD is the glomerular filtration rate (GFR), which significantly declines in each stage. Renal failure is also identified using the GFR. If  $GFR < 15$  ml/min, this means that the kidneys have failed or are about to fail. Chronic renal disease has reached its end stage. Medically speaking, It is challenging to diagnose CKD because it has a wide range of symptoms. Clinical judgment largely depends on the expertise and experience of the physician examining the symptoms of the patient. As the healthcare industry develops and new drugs become available, it becomes increasingly difficult for physicians to keep up with current changes in clinical practice.

In this regard, machine learning classification algorithms play a vital part. Using various machine learning algorithms, the prognosis rate of the disease can be studied, and evaluated and proper actions can be taken based on the results and discussions. Machine learning models can effectively aid clinicians to achieve this goal due to their fast and accurate recognition performance. Using machine learning algorithms for medical studies, the disease can be predicted with a high accuracy rate and a very short time. Machine learning techniques have become reliable for medical treatment. In this project, we propose a machine-learning methodology for diagnosing CKD. Our project aims to predict Chronic Kidney Disease based on full features and important features of the CKD dataset.



**Fig 1.1 Theoretical flow of Chronic Kidney Disease**

## **1.1 SCOPE**

The benefit of this approach is that the prediction process takes far less time and helps doctors to initiate treatment at the earliest for patients with CKD and further to classify larger populations of patients within a shorter span. Since the project is Machine learning based, the cost spent in executing this project would not demand a cost for software and related products, as most of the products are open source and free to use. Hence the project would consume minimal cost and is economically feasible. This project would help detect the chances of a person having CKD further on in his life which would be really helpful and cost-effective for people. Patients would not have to go to a doctor unless they are flagged by the algorithms. This would make it cheaper and easier for the modern busy person.

## **1.2 SOCIETAL / ENVIRONMENTAL IMPACT / NOVELTY OF IDEA**

- Creating a web application where a user can log in and upload the dataset.
- Check if he/she is diagnosed with CKD based on ML algorithms.
- CKD remains undetected in its early stage and the patients can only realize the severity of the disease when it gets advanced. Hence, detecting such disease at an earlier stage is a key challenge now.

## **CHAPTER 2**

### **PROBLEM DEFINITION**

## **CHAPTER 2 PROBLEM DEFINITION**

In the last 15 years' data, it has been noticed that there is an increase in the number of patients who are suffering from CKD disease, and more than 60% of patients are not receiving medical attention. CKD ranks 27 and 18 in 1990 and 2010 respectively as the world's prime reason for death and is expected to remain the most prevalent stage in 2020 and 2030. 956,000 people died in 2013 because of CKD. At the last stage, the patient must take dialysis or kidney transplantation. One of the best ways to reduce this death rate is early treatment.

But in developing countries, patients take medicine when they reach a severe state. An automated system can be built to detect CKD-affected patients before reaching the last stage. We aim to compute, analyze and compare Machine Learning classification approaches to determine which classification approach is optimal for the prediction of chronic kidney disease.

# **CHAPTER 3**

## **LITERATURE REVIEW**



## CHAPTER 3 LITERATURE REVIEW

Machine learning algorithms have become a very essential tool in not just the field of math and engineering but medicine as well. Using machine learning algorithms, predictions are done on a regular basis in order to take quick and necessary actions. Mostly, to improve performance and accuracy, algorithms are implemented.

In [1] research work, authors used artificial neural networks for machine learning and deep neural network-based analysis. The results were computed based on full features, and selected features by CFS. To test the performance of machine learning classifier algorithms, one deep neural network model was built and the results were noted. In some cases, the deep neural network gave strong results and important features were extracted by itself, that is, no feature selection algorithm was required. The authors mentioned that Logistic and KNN classifiers give satisfactory results and have negligible differences between precision and recall values. In comparison with them, precision for Logistic and KNN classifiers is low whereas recall is high. The results were computed based on full features, and selected features by CFS. we infer that Logistics and KNN did not give suitable results. SMOTE(Synthetic Minority Oversampling Technique) gave better results with selected features by Least Absolute Shrinkage Selection Operator regression.

In [2] authors works with the theoretical capacity of research work where authors studied human activities, Using the data acquired from the UCI machine learning repository, and have used four different machine learning algorithm such as Logistic Regression algorithm, Decision Tree algorithm, Random Forest algorithm, K-Nearest Neighbors algorithm are implemented in the pre-processed data in order to get the accurate prediction rate for the dataset collected. These models successfully generated the best accuracy rate on the recognized dataset except the KNN algorithm with a higher error rate comparatively. The authors also give the conclusion that the Random forest takes more time to predict results and the best rating in the ROC curve. Hence, an accurate rate of prediction is undoubtedly dependent on the pre-processed technique, the techniques of the pre-processed model must be handled carefully to achieve more correct results.

In [3] authors have aimed to diagnose Chronic Kidney Disease (CKD) at an earlier stage, this manuscript introduced a variety of machine learning algorithms. The models obtained from CKD

patients are trained and authenticated with the mentioned input parameters in the paper. Support Vector Machine, Logistic Regression, and KNN are analyzed to conduct the study of CKD. The performances of those algorithms were determined primarily on the basis of precision. The authors' results exemplified that the Support Vector Machine algorithm predicts Chronic Kidney Disease better than Logistic Regression and K-Nearest Neighbors. The accuracy for SVM is 99.25, Logistic Regression is 77.25 and K-Nearest Neighbors is 78.75. The benefit of this approach is that the prediction process takes far less time and helps doctors to initiate treatment at the earliest for patients with CKD and further to classify larger populations of patients within a shorter span. In addition, we have inferred that to help minimize the incidence of CKD, it has been an attempt to predict if a person with this disease has the chronic risk factors are hypertension, family history of kidney failure, and diabetes using the appropriate dataset.

[4] This work proposed a workflow to predict CKD status based on clinical data, incorporating data preprocessing, a missing value handling method with collaborative filtering and attribute selection. Out of the 11 machine learning methods considered, the extra tree classifier and random forest classifier are shown to result in the highest accuracy and minimal bias to the attributes. The authors also considered the practical aspects of data collection and highlighted the importance of incorporating domain knowledge when using machine learning for CKD status prediction. The proposed methodology consists of 3 key steps: Data preprocessing, model training, and model selection.

The data distribution has properly covered the whole domain in CKD, but the general attributes like appetite, anemia, and pedal edema are biased towards CKD. It is easy to achieve an accurate prediction using this data set but in the general context, it may lead to false positives. Further, the missing values made it impossible to achieve perfect accuracy without filling them from a collaborative imputer instead of a constant. Considering the medical importance of the attributes, some of them have a lesser correlation compared to others because of the stage they appear in the patient. When training the models makes a huge impact on the accuracy.

In order to acquire an accurate predictive rate for the dataset that is presented in this paper [5], the authors employed four distinct machine-learning algorithms. When all of the introduced strategies

were contrasted, the effective results were acquired from four ideal machine learning techniques, namely Logistic Regression and Random Forest. On the recognized dataset, These models are capable of producing 100% accuracy. However, among the other exhibited classifiers, K received the lowest predictable score (95.83 percent), which was quite low. The best ROC curve rating and the longer time it takes to predict results are also evident. In addition, the KNN algorithm has a higher error rate than other algorithms. The pre-processed model's techniques need to be handled with care in order to produce more accurate results because the pre-processed technique unquestionably affects the rate of accurate prediction.

This research [6] developed a strategy that combines a cost-sensitive AdaBoost classifier with a Selection of features depending on information acquisition to enhance the detection of chronic renal disease. As a starting point for comparison, the performance of six additional machine learning classifiers was used. The classifiers are SVM, LR, DT, RF, XGBoost, and AdaBoost. First, the importance of various traits was determined and ranked using the IG approach. Second, the classifiers were trained using both full feature sets and modified feature sets. The outcome of the experiment demonstrates that some attributes improved the effectiveness of the classification. Additionally, the suggested AdaBoost fared better than other classifiers and techniques in recent research. As a result, the combination of the IG-based feature extraction approach with the AdaBoost method is an effective method for identifying CKD.

In this work [7], CKD was predicted using a mix of feature extraction approaches and classification machine learning algorithms developed on big data platforms (Apache Spark). The key features from the dataset were chosen using the Relief-F and chi-squared feature extraction algorithms. Machine learning algorithms such as Decision Tree, Logistic Regression, Random forest, Naive Bayes, Support vector Machine, and Gradient Boosting Classifier were used on a benchmark dataset for chronic renal disease. They were implemented on both the entire set of characteristics and the features that were selected. Error Metrics were also used to register the cross-validation results and the test dataset. To improve the ML characteristics, grid search & cross-validation were used. Support vector Machine (SVM), Gradient Boosting (GD), and Decision Tree Classifiers(DT) using the specified features performed the best, according to the results.

The number of individuals requiring kidney transplants as a result of chronic renal disease is rising daily [8]. With a mortality rate of 5.44%, CKD ranks as the fourth most common cause of death in Saudi Arabia. To identify CKD, ML techniques were initially tested on the medical records of Saudi Arabia. To choose features, the authors utilized recursive feature elimination and correlation coefficients. Then, four classification algorithms— SVM, ANN, KNN, and Naive Bayes, were investigated. The precision, recall, accuracy, and f-measure of each of these classifiers were assessed in order to determine their efficacy. ANN, SVM, and NB all reached 98% accuracy, whereas k-NN only managed 93.9%.

## **CHAPTER 4**

### **PROJECT DESCRIPTION**

## CHAPTER 4 PROJECT DESCRIPTION

### 4.1 PROPOSED DESIGN

KNN imputation is used to fill in the missing values. To our knowledge, this is the first time that KNN imputation has been used for the diagnosis of CKD. In addition, building an integrated model is also a good way to improve the performance of separate individual models. The proposed measurements before being diagnosed.

In addition, the resulting integrated model shows a higher accuracy. Therefore, it is speculated that this methodology might be applicable to the clinical data in the actual medical diagnosis.

#### 4.1.1 ARCHITECTURE

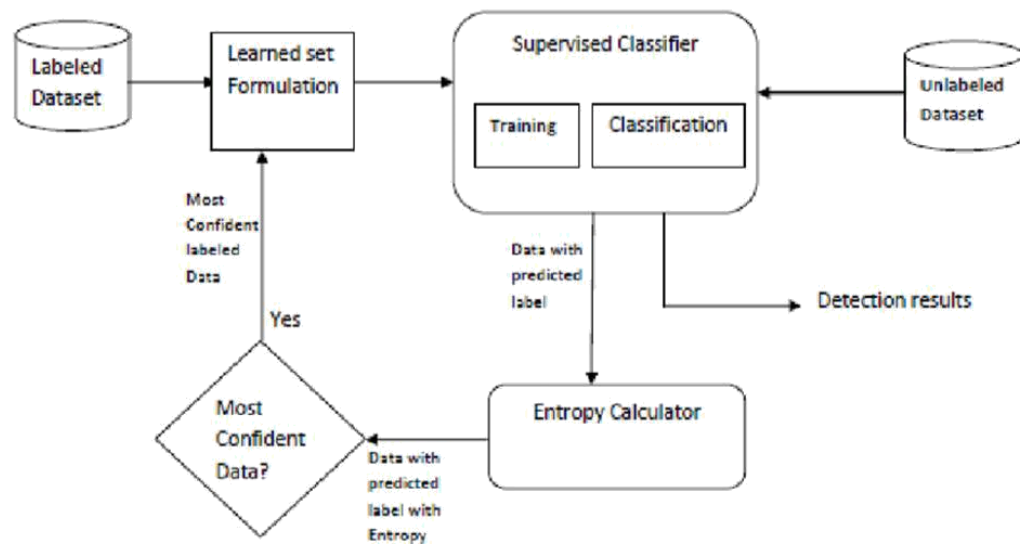


Fig 4.1.1 Architecture Design

## **CHAPTER 5**

### **REQUIREMENTS**

## **CHAPTER 5 REQUIREMENTS**

### **5.1 FUNCTIONAL REQUIREMENTS**

- The user must be able to log in and upload the dataset.
- All the data must be in the same format as structured data.
- The data collected will be vectorized and sent across to the classifier.
- Create a model to predict if a person has CKD.
- Users should be able to enter the parameter values which are used to determine if a person has CKD or not.

### **5.2 NON - FUNCTIONAL REQUIREMENTS**

- Organizational Requirements: Clinical Laboratory, Medical Laboratory
- Usability: easy to use for even a non-technical user.
- Security: admin roles can log in and can control data.

### **5.3 HARDWARE AND SOFTWARE REQUIREMENTS**

- Operating System : Windows 11
- Memory :8GBRAM
- Free Space : 65 GB of FreeSpace
- IDLE : Python Platform
- Language : Python
- CPU : Core i5, Ryzen 5
- Technology : Machine Learning
- Screen Resolution : 1280x1024 or larger
- Application Window Size : 1024x680 or larger



## **CHAPTER 6**

### **METHODOLOGY**

# CHAPTER 6 METHODOLOGY

## 6.1 MODULES

- Data Collection
- Dataset
- Data Preparation
- Model Selection
- Analyze and Prediction
- Accuracy on test set
- Saving the Trained Model

### 6.1.1 DATA COLLECTION

This is the first real step towards the real development of a machine learning model, collecting data. This is a critical step that will cascade in how good the model will be, the more and better data that we get, the better our model will perform.

There are several techniques to collect the data, like web scraping, manual interventions and etc.

The dataset used in this Chronic kidney disease dataset taken from UCI:[https://archive.ics.uci.edu/ml/datasets/chronic\\_kidney\\_disease](https://archive.ics.uci.edu/ml/datasets/chronic_kidney_disease)

## 6.1.2 DATASET DESCRIPTION

The dataset consists of 400 individual data. There are 26 columns in the dataset, which are described below.

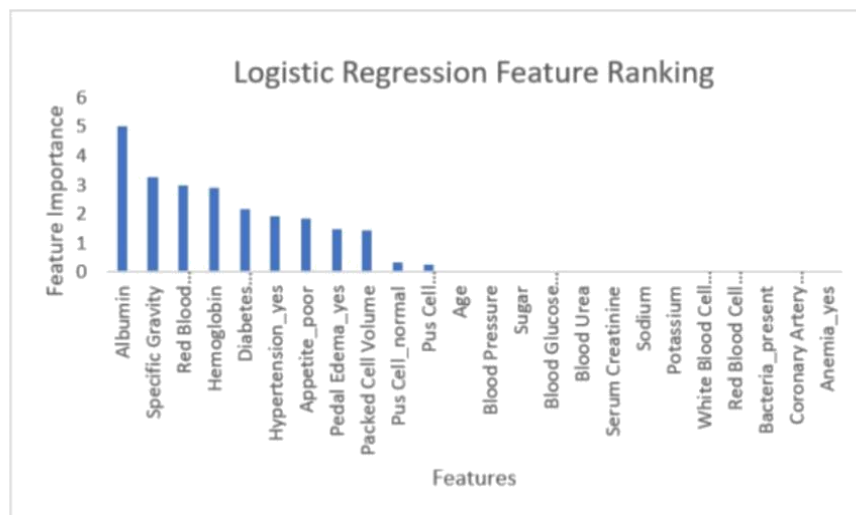
age	-	age
bp	-	blood pressure
sg	-	specific gravity
al	-	albumin
su	-	sugar
rbc	-	red blood cells
pc	-	pus cell
pcc	-	pus cell clumps
ba	-	bacteria
bgr	-	blood glucose random
bu	-	blood urea
sc	-	serum creatinine
sod	-	sodium
pot	-	potassium
hemo	-	hemoglobin
pcv	-	packed cell volume
wc	-	white blood cell count
rc	-	red blood cell count
htn	-	hypertension
dm	-	diabetes mellitus
cad	-	coronary artery disease
appet	-	appetite
pe	-	pedal edema
ane	-	anemia
class	-	classification

## 6.1.3 DATA PREPARATION

We will transform the data. By getting rid of missing data and removing some columns. First, we will create a list of column names that we want to keep or retain. Next, we drop or remove all columns except for the columns that we want to retain. Finally, we drop or remove the rows that have missing values from the data set. Split into training and evaluation sets.

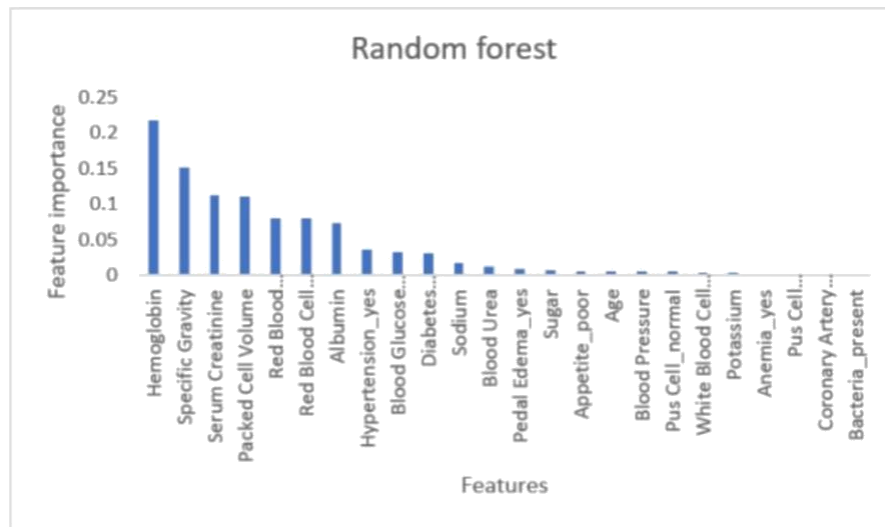
## 6.1.4 MODEL SELECTION

It is a supervised learning algorithm that includes more dependent variables. Linear regression analysis demands that the dependent variable is continuous. For example, a patient may or may not be affected by a given disease, or he can die or survive during a given time period. Logistic regression analysis is a statistical technique that describes the relationship between an independent variable (either continuous or not) and a dependent variable (or dummy variable) (that is, a variable with only two possible values: 0=outcome absent and 1=outcome present).



**Fig 6.1.4.1 Logistic Regression**

The random forest starts with a standard machine learning technique called a ‘decision tree’ which, in ensemble terms, corresponds to our weak learner. The decision tree algorithm repeatedly splits the data set according to a criterion that maximizes the separation of the data, resulting in a tree-like structure. In this algorithm, an input is entered at the top and as it traverses down the tree the data gets bucketed into smaller and smaller sets. The random forest takes this notion to the next level by combining trees with the notion of an ensemble.



**Fig 6.1.4.2 Random Forest**

### 6.1.3 ANALYZE AND PREDICTION

In the actual dataset, we chose only 19 features:

1. Age(numerical) --> age in years
2. Blood Pressure(numerical) bp in mm/Hg
3. Specific Gravity(nominal) sg - (1.005,1.010,1.015,1.020,1.025)
4. Albumin(nominal)al - (0,1,2,3,4,5)
5. Sugar(nominal) su - (0,1,2,3,4,5)
6. Blood Glucose Random(numerical) bgr in mgs/dl
7. Blood Urea(numerical) bu in mgs/dl
8. Serum Creatinine(numerical) sc in mgs/dl
9. Sodium(numerical) sod in mEq/L
10. Potassium(numerical) pot in mEq/L
11. Hemoglobin(numerical) hemo in gms
12. Packed Cell Volume(numerical)
13. White Blood Cell Count(numerical) wc in cells/cumm
14. Hypertension(nominal) htn - (yes,no)
15. Diabetes Mellitus(nominal) dm - (yes,no)
16. Coronary Artery Disease(nominal) cad - (yes,no)
17. Appetite(nominal) ppet - (good,poor)
18. Pedal Edema(nominal) pe - (yes,no)
19. Anemia(nominal)ane - (yes,no)
20. Chronic\_kidney\_disease: Displays whether the individual is suffering from kidney disease or no

# **CHAPTER 7**

## **EXPERIMENTATION**

# CHAPTER 7 EXPERIMENTATION

## 7.1 AGE VS ALL CONTINUOUS COLUMNS

The below diagram is the scatter plot for the age Vs all continuous columns (Numericals), to identify the relation between the numerical.

```
plt.figure(figsize=(20,15), facecolor='white')
plotnumber = 1

for column in contcols:
    if plotnumber<=11 :      # as there are 11 continous columns in the data
        ax = plt.subplot(3,4,plotnumber) # 3,4 is refer to 3X4 matrix
        plt.scatter(data['age'],data[column]) #plotting scatter plot
        plt.xlabel(column,fontsize=20)
        #plt.ylabel('Salary',fontsize=20)
        plotnumber+=1
plt.show()
```

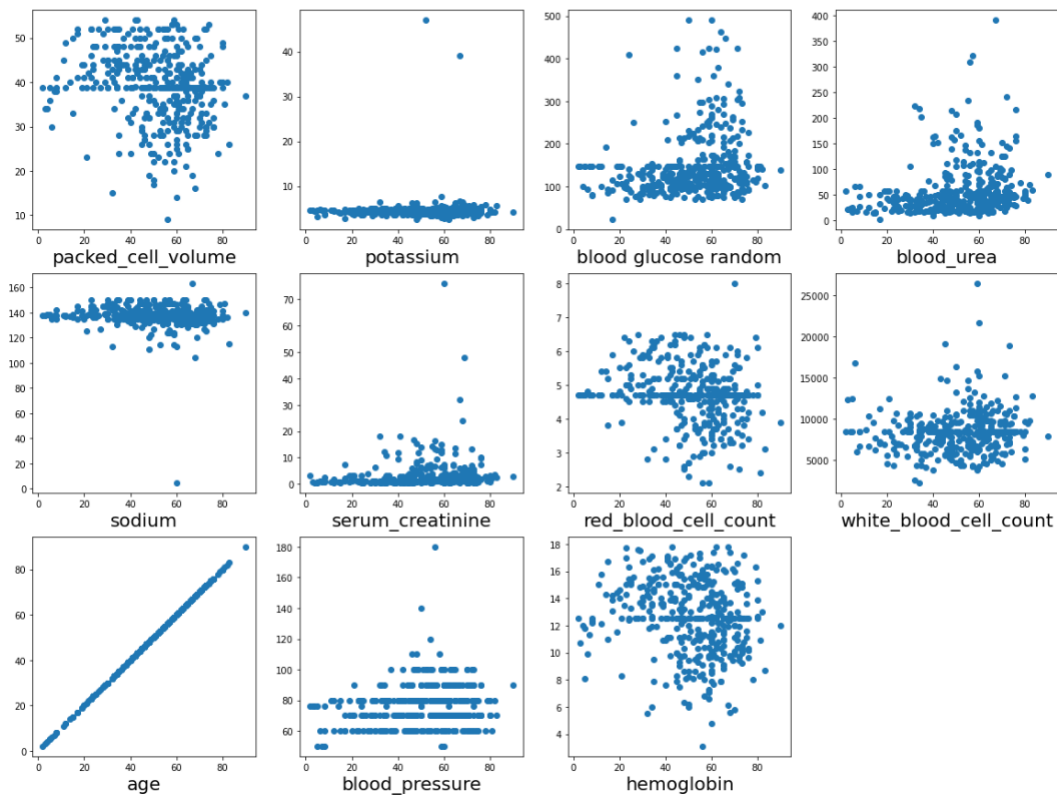


Fig 7.1 Age vs Continuous Columns



Heatmaps show relationships between two variables, one plotted on each axis. By observing how cell colors change across each axis, you can follow if there are any patterns in value for one or both variables. It shows the correlation between dependent and independent variables.

age	1.00	0.15	-0.16	0.09	0.19	-0.02	-0.10	0.15	0.04	0.21	0.19	0.13	-0.09	0.05	-0.18	-0.22	0.10	-0.20	0.40	0.36	0.23	0.15	0.10	0.06	-0.23
	0.15	1.00	-0.16	0.12	0.19	-0.15	-0.16	0.06	0.11	0.15	0.18	0.14	-0.10	0.07	-0.28	-0.29	0.03	-0.22	0.27	0.23	0.09	0.18	0.05	0.19	-0.29
blood pressure	-0.16	-0.16	1.00	-0.48	-0.29	0.25	0.37	-0.31	-0.23	-0.32	-0.25	-0.18	0.22	-0.06	0.49	0.50	-0.21	0.44	-0.32	-0.35	-0.14	-0.23	-0.25	-0.18	0.66
	0.09	0.12	-0.48	1.00	0.29	-0.39	-0.56	0.42	0.38	0.31	0.35	0.16	-0.23	0.11	-0.47	-0.48	0.21	-0.41	0.41	0.31	0.20	0.30	0.41	0.23	0.53
specific gravity	0.19	0.19	-0.29	0.29	1.00	-0.09	-0.19	0.17	0.12	0.63	0.13	0.09	-0.05	0.18	-0.16	-0.18	0.16	-0.16	0.25	0.43	0.23	0.07	0.12	0.04	-0.29
	-0.02	-0.15	0.25	-0.39	-0.09	1.00	0.38	0.10	-0.18	-0.15	-0.24	-0.14	0.14	0.02	0.28	0.28	-0.00	0.20	-0.14	-0.15	-0.11	0.16	-0.20	-0.11	0.28
albumin	-0.10	-0.16	0.37	-0.56	-0.19	0.38	1.00	-0.52	-0.33	-0.26	-0.34	-0.16	0.17	-0.16	0.41	0.42	-0.11	0.38	-0.29	-0.20	-0.17	-0.27	0.35	-0.26	0.38
	0.15	0.06	-0.31	0.42	0.17	-0.10	-0.52	1.00	0.28	0.20	0.18	0.05	-0.14	-0.01	-0.28	-0.29	0.16	-0.27	0.20	0.17	0.19	0.19	0.10	0.18	-0.27
red blood cells	0.04	0.11	-0.23	0.38	0.12	-0.18	-0.33	0.28	1.00	0.09	0.16	0.05	-0.08	-0.00	-0.20	-0.19	0.10	-0.19	0.09	0.08	0.16	0.15	0.13	0.05	-0.19
	0.21	0.15	-0.32	0.31	0.63	-0.15	-0.26	0.20	0.09	1.00	0.13	0.08	-0.15	0.06	-0.27	-0.27	0.12	-0.22	0.37	0.50	0.21	0.18	0.10	0.13	-0.40
pus_cell_clumps	0.19	0.18	-0.25	0.35	0.13	-0.24	-0.34	0.18	0.16	0.13	1.00	0.58	-0.31	0.34	-0.54	-0.53	0.04	-0.47	0.39	0.31	0.22	0.27	0.34	0.44	0.37
	0.13	0.14	-0.18	0.16	0.09	-0.14	-0.16	0.05	0.05	0.08	0.58	1.00	-0.62	0.21	-0.34	-0.34	-0.01	-0.32	0.27	0.21	0.19	0.16	0.18	0.24	-0.29
blood glucose random	-0.09	-0.10	0.22	-0.23	0.05	0.14	0.17	-0.14	-0.08	-0.15	-0.31	-0.62	1.00	0.07	0.33	0.35	0.01	0.32	-0.31	-0.27	-0.22	-0.16	-0.15	-0.20	0.34
	0.05	0.07	-0.06	0.11	0.18	0.02	-0.16	-0.01	-0.00	0.06	0.34	0.21	0.07	1.00	-0.10	-0.12	-0.07	0.12	0.06	0.06	0.01	-0.02	0.06	0.10	-0.08
serum creatinine	-0.18	-0.28	0.49	-0.47	-0.16	0.28	0.41	-0.28	-0.20	-0.27	-0.54	-0.34	0.33	-0.10	1.00	0.85	-0.15	0.68	-0.58	-0.47	-0.28	-0.39	-0.38	0.56	0.73
	-0.22	-0.29	0.50	-0.48	-0.18	0.28	0.42	-0.29	-0.19	-0.27	-0.53	-0.34	0.35	-0.12	0.85	1.00	-0.18	0.70	-0.57	-0.45	-0.29	-0.39	-0.39	0.51	0.69
hemoglobin	0.10	0.03	-0.21	0.21	0.16	-0.00	-0.11	0.16	0.10	0.12	0.04	-0.01	0.01	-0.07	-0.15	-0.18	1.00	-0.15	0.12	0.15	0.01	0.15	0.14	0.04	0.21
	-0.20	-0.22	0.44	-0.41	-0.16	0.20	0.38	-0.27	-0.19	-0.22	-0.47	-0.3													

24

## 7.3 SPLITTING THE DATASET

Splitting the data set into Two, the Train dataset(X\_train,y\_train) and Test dataset(X\_test,y\_test).

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=2)#train test split
print(x_train.shape)
print(y_train.shape)
print(x_test.shape)
print(y_test.shape)
```

Fig 7.3 Splitting the data into Train and Test

## 7.4 MACHINE LEARNING MODEL

```
from sklearn.linear_model import LogisticRegression
lgr = LogisticRegression()
lgr.fit(x_train,y_train)

y_pred = lgr.predict([[129,99,1,0,0,1,0,1]])

print(y_pred)
c(y_pred)
```

Fig 7.4 Building a Machine Learning Model

## **CHAPTER 8**

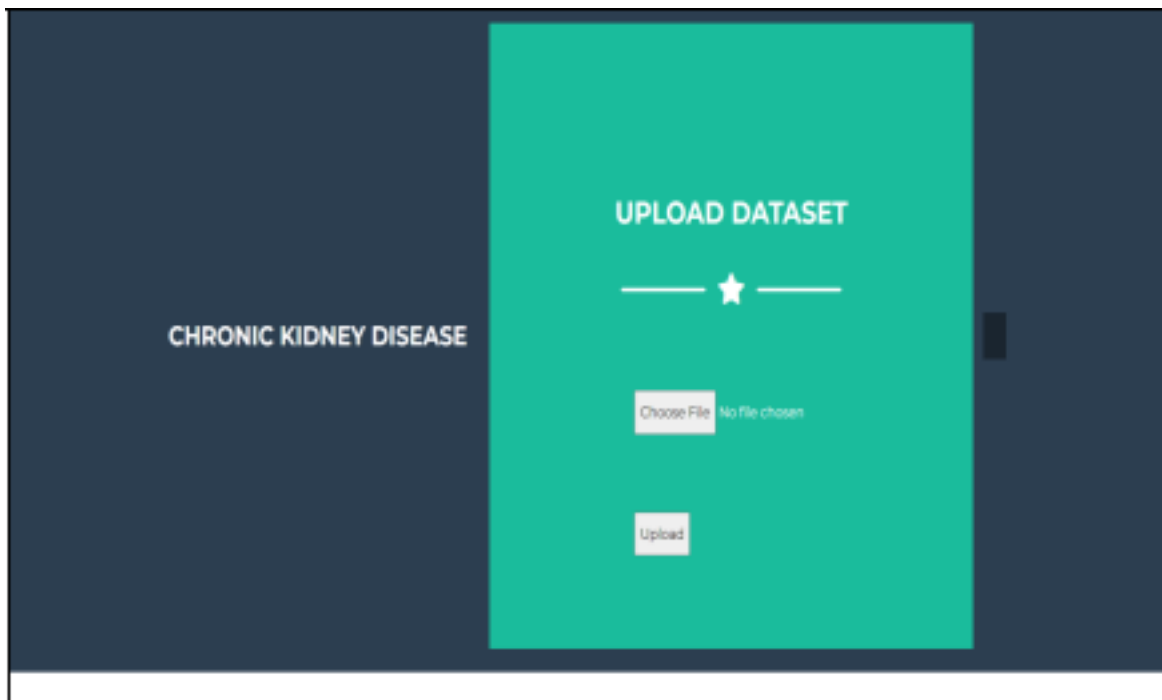
### **TESTING AND RESULTS**

## CHAPTER 8 TESTING AND RESULTS

### 8.1 RESULT (INTERFACE)

#### 8.1.1 DATASET UPLOAD PAGE

An Interface starting with a home page includes uploading a dataset as one of its features. Users can choose a file and upload a dataset to train the model.



**Fig 8.1.1 Dataset Upload Page**

## 8.1.2 PREVIEW OF DATASET

This is the page where a preview of the dataset which was uploaded is displayed before training. Next, we can train the dataset as shown in Fig.8.1.2.2

CHRONIC KIDNEY DISEASE													
PREVIEW													
★													
	age	bp	sg	al	su	bgr	bu	sc	sod	pot	hemo	pcv	
Id													
1	48.000000	80.000000	1.020000	1.000000	0.000000	121.000000	36.000000	1.200000	135.297553	3.000617	15.400000	44.000000	7800.000000
2	7.000000	50.000000	1.020000	4.000000	0.000000	134.383588	18.000000	0.800000	131.564702	4.401237	11.300000	38.000000	6000.000000
3	62.000000	80.000000	1.010000	2.000000	3.000000	423.000000	53.000000	1.800000	134.149877	6.809881	9.600000	31.000000	7500.000000
4	48.000000	70.000000	1.005000	4.000000	0.000000	117.000000	56.000000	3.800000	111.000000	2.500000	11.200000	32.000000	6700.000000

Fig 8.1.2.1 Preview of Dataset

CHRONIC KIDNEY DISEASE													
395	50.000000	80.000000	1.020000	0.000000	0.000000	137.000000	46.000000	0.800000	139.000000	5.000000	14.100000	45.000000	9500.000000
396	55.000000	80.000000	1.020000	0.000000	0.000000	140.000000	49.000000	0.500000	150.000000	4.900000	15.700000	47.000000	6700.000000
397	42.000000	70.000000	1.025000	0.000000	0.000000	75.000000	31.000000	1.200000	141.000000	3.500000	16.500000	54.000000	7800.000000
398	12.000000	80.000000	1.020000	0.000000	0.000000	100.000000	26.000000	0.600000	137.000000	4.400000	15.800000	49.000000	6600.000000
399	17.000000	60.000000	1.025000	0.000000	0.000000	114.000000	50.000000	1.000000	135.000000	4.900000	14.200000	51.000000	7200.000000
400	58.000000	80.000000	1.025000	0.000000	0.000000	131.000000	18.000000	1.100000	141.000000	3.500000	15.800000	53.000000	6800.000000
Click to Train   Test													

Fig 8.1.2.2 End of Dataset And Train

## 8.2 TEST CASES

### 8.2.1 TEST CASE 1: PATIENT SUFFERING WITH CKD

From the test data, input is given for the features. This is the data that shows that the patient is suffering from CKD.

Feature	Value
Age	121
Blood Pressure (mmHg)	36
Blood Sugar (mg/dl)	1.2
Cholesterol (mg/dl)	135.2976
Glucose (mg/dl)	3.000617
Hemoglobin (g/dl)	15.4
Platelets (thousands/mm <sup>3</sup> )	44
Red Blood Cells (millions/mm <sup>3</sup> )	7800
Serum Creatinine (mg/dl)	1
Serum Urea Nitrogen (mg/dl)	1
Serum Bilirubin (mg/dl)	0
Serum Albumin (g/dl)	1
Serum Calcium (mg/dl)	0
Serum Potassium (mEq/L)	0
Serum Sodium (mEq/L)	0

Predict

The Patient has Chronic Kidney Disease Analysis

Fig 8.2.1 Patient with CKD

## 8.2.2 TEST CASE 2: PATIENT WITHOUT CKD

This is the data that shows the patient is not suffering from CKD i.e. Normal.

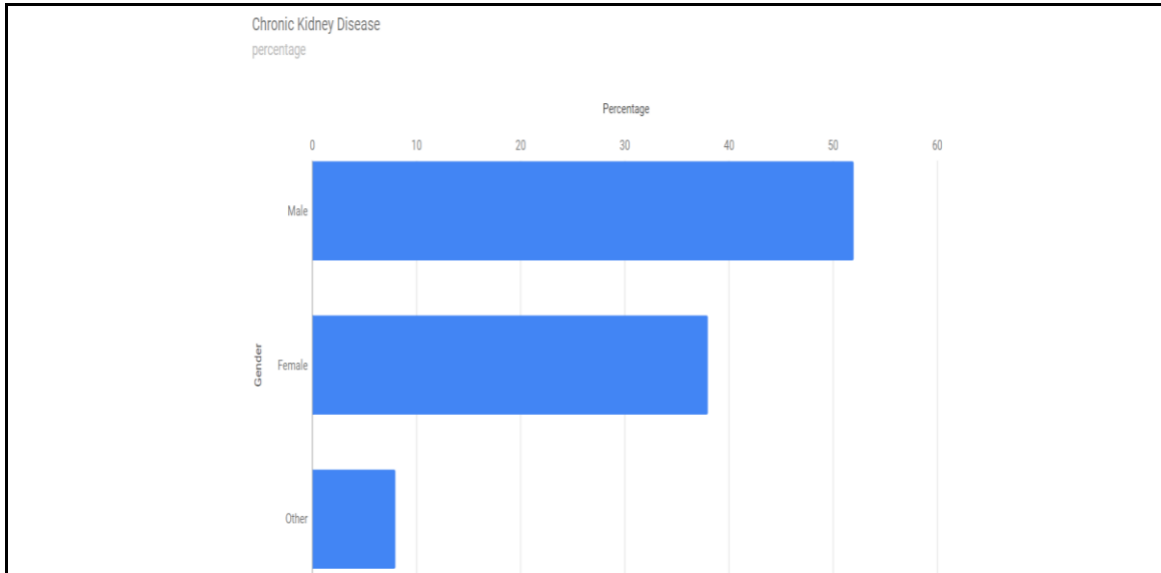
The screenshot shows a web application titled "Kidney Disease Diagnosis". On the right side, there are 14 input fields containing the following values: 114, 50, 1, 135, 4.9, 14.2, 51, 7200, 0, 0, 0, 1, 0, and d. Below these fields is a blue "Predict" button. On the left side, the text "The Patient has Normal" is displayed, followed by a blue "Analysis" button.

**Fig 8.2.2 Patient without CKD(Normal)**

## 8.3 ANALYSIS OF DATASET

### 8.3.1 BAR GRAPH

Dataset uploaded can be analyzed in the form of BarGraph. From this bar Graph, it is analyzed that the number of males suffering from CKD is 50% whereas females suffering is between 30-40%.

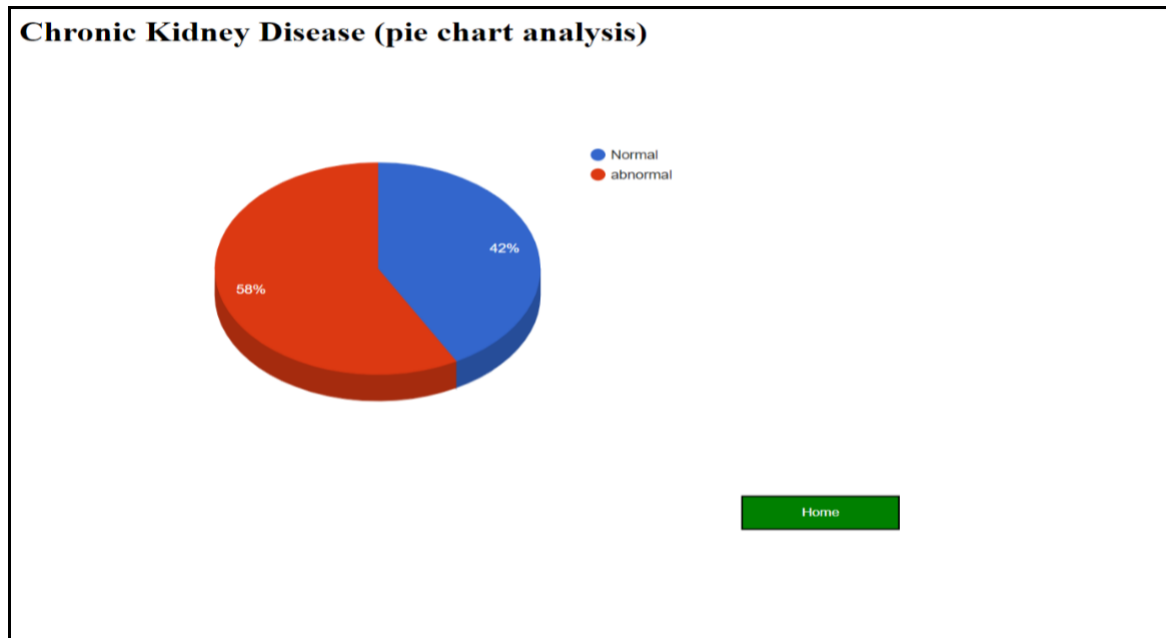


**Fig 8.3.1 Analysis Through Bar Graph**



### 8.3.2 PIE CHART

From the pie Chart we analyzed from the dataset that 42% are normal people who are not suffering from CKD whereas 58% are abnormal i.e. suffering from CKD.



**Fig 8.3.2 Analysis through Pie Chart**

## **CHAPTER 9**

### **CONCLUSION AND FUTURE WORK**

## **9.1 CONCLUSION**

The proposed CKD diagnostic methodology is feasible in terms of data imputation and sample diagnosis. After the unsupervised imputation of missing values in the data set by KNN imputation, the integrated model achieved satisfactory accuracy. Hence, we speculate that applying this methodology to the practical diagnosis of CKD would achieve a desirable effect. In addition, this methodology might be applicable to the clinical data of other diseases in actual medical diagnosis. However, in the process of establishing the model, due to the limitations of the conditions, the available data samples are relatively small, including only 400 samples. Therefore, the generalization performance of the model might be limited. In addition, due to there being only two categories (CKD and not CKD) of data samples in the data set, the model can not diagnose the severity of CKD.

## **9.2 SCOPE FOR FUTURE WORK**

In the future, a large number of more complex and representative data will be collected to train the model to improve the generalization performance while enabling it to detect the severity of the disease. We believe that this model will be more and more perfect with the increase in size and quality of the data.

## REFERENCES

- [1] Chittora, P., Chaurasia, S., Chakrabarti, P., Kumawat, G., Chakrabarti, T., Leonowicz, Z., ... & Bolshev, V. (2021). Prediction of chronic kidney disease-a machine learning perspective. *IEEE Access*, 9, 17312-17334.
- [2] Shamrat, F. J. M., Ghosh, P., Sadek, M. H., Kazi, M. A., & Shultana, S. (2020, November). Implementation of machine learning algorithms to detect the prognosis rate of kidney disease. In *2020 IEEE International Conference for Innovation in Technology (INOCON)* (pp. 1-7). IEEE.
- [3] Gudeti, B., Mishra, S., Malik, S., Fernandez, T. F., Tyagi, A. K., & Kumari, S. (2020, November). A novel approach to predict chronic kidney disease using machine learning algorithms. In *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 1630-1635). IEEE.
- [4] Ekanayake, I. U., & Herath, D. (2020, July). Chronic kidney disease prediction using machine learning methods. In *2020 Moratuwa Engineering Research Conference (MERCon)* (pp. 260-265). IEEE.
- [5] Shamrat, F. J. M., Ghosh, P., Sadek, M. H., Kazi, M. A., & Shultana, S. (2020, November). Implementation of machine learning algorithms to detect the prognosis rate of kidney disease. In *2020 IEEE International Conference for Innovation in Technology (INOCON)* (pp. 1-7). IEEE.
- [6] Ebiaredoh-Mienye, S. A., Swart, T. G., Esenogho, E., & Mienye, I. D. (2022). A machine learning method with filter-based feature selection for improved prediction of chronic kidney disease. *Bioengineering*, 9(8), 350.
- [7] Abdel-Fattah, M. A., Othman, N. A., & Goher, N. (2022). Predicting Chronic Kidney Disease Using Hybrid Machine Learning Based on Apache Spark. *Computational Intelligence and Neuroscience*, 2022.
- [8] Alassaf, R. A., Alsulaim, K. A., Alroomi, N. Y., Alsharif, N. S., Aljubeir, M. F., Olatunji, S. O., ... & Alturayef, N. S. (2018, November). Preemptive diagnosis of chronic kidney disease using machine learning techniques. In *2018 international conference on innovations in information technology (IIT)* (pp. 99-104). IEEE.

- [9] Hossain, M. M., Swarna, R. A., Mostafiz, R., Shaha, P., Pinky, L. Y., Rahman, M. M., ... & Iqbal, M. S. (2022). Analysis of the performance of feature optimization techniques for the diagnosis of machine learning-based chronic kidney disease. *Machine Learning with Applications*, 100330.
- [10] Ma, F., Sun, T., Liu, L., & Jing, H. (2020). Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural networks. *Future Generation Computer Systems*, 111, 17-26.
- [11] Antony, L., Azam, S., Ignatious, E., Quadir, R., Beeravolu, A. R., Jonkman, M., & De Boer, F. (2021). A comprehensive unsupervised framework for chronic kidney disease prediction. *IEEE Access*, 9, 126481-126501.
- [7] Antony, L., Azam, S., Ignatious, E., Quadir, R., Beeravolu, A. R., Jonkman, M., & De Boer, F. (2021). A comprehensive unsupervised framework for chronic kidney disease prediction. *IEEE Access*, 9, 126481-126501.

## **APPENDIX**

# SAMPLE CODE

## App.py

```
import numpy as np
from flask import Flask, jsonify, request, render_template
from flask_cors import CORS, cross_origin
import pickle
import pandas as pd

app = Flask(__name__)
cors = CORS(app)
app.config['CORS_HEADERS'] = 'Content-Type'
model = pickle.load(open('kid1.pkl', 'rb'))
#json_file = open('model.json', 'r')

#loaded_model_json = json_file.read()
#json_file.close()

#loaded_model = model_from_json(loaded_model_json)

#loaded_model.load_weights("model9954.h5")
#print("Loaded model from disk")

@app.route('/')
@app.route('/first')
def first():
    return render_template('first.html')
@app.route('/index')
def index():
    return render_template('index.html')
@app.route('/abstract')
def abstract():
    return render_template('abstract.html')
@app.route('/future')
def future():
    return render_template('future.html')

@app.route('/chart')
def chart():
    return render_template('chart.html')
@app.route('/pie')
def pie():
    return render_template('pie.html')
@app.route('/upload')
def upload():
    return render_template('upload.html')
@app.route('/preview', methods=['POST'])
def preview():
    if request.method == 'POST':
        dataset = request.files['datasetfile']
        df = pd.read_csv(dataset, encoding = 'unicode_escape')
        df.set_index('Id', inplace=True)
        return render_template("preview.html", df_view = df)
```

```

@app.route('/home')
def home():
    return render_template('index2.html')

@app.route('/predict',methods=['POST'])
@cross_origin()
def predict():
    """
    For rendering results on HTML GUI
    """

    # if("text" == "M"):
    #     text = 0
    # else:
    #     text = 1

    int_features = [float(x) for x in request.form.values()]
    final_features = [np.array(int_features)]

    # prediction = model.predict_proba(final_features)
    prediction = model.predict(final_features)

    output = prediction[0]
    if output<0.5:
        output1='Normal'
    elif output>0.5:
        output1='Chronic Kidney Disease'

    return render_template('index2.html', prediction_text='The Patient has {}'.format(output1))

if __name__ == "__main__":
    app.run(debug=True)

```

## First.html

```

<!DOCTYPE html>
<html lang="en">

<head>
<meta charset="UTF-8" />
<meta name="viewport" content="width=device-width, initial-scale=1.0" />
<link rel="stylesheet" href="../static/css/estilos.css" /> <link
href="https://fonts.googleapis.com/css2?family=Bebas+Neue&family=Open+Sans:wght@400;600&display=swap"
rel="stylesheet" />

<title> Chronic Kidney Disease</title>
</head>

<body>

```



```

<header>
  <div class="contenedor">
    <h2 class="logotipo"> Chronic Kidney Disease</h2>
    <nav>
      <a href="{{ url_for('index') }}">Login </a>
      <a href="{{ url_for('abstract') }}">Abstract</a>
      <a href="{{ url_for('upload') }}">Upload Data </a>
      <a href="{{ url_for('chart') }}">Chart</a>
      <a href="{{ url_for('future') }}">Future</a>
    </nav>
  </div>
</header>
<body>
<main>
  <div class="pelicula-principal">
    <div class="contenedor">
      <h3 class="titulo">Chronic Kidney Disease Prediction Using ML</h3>

    </div>
  </div>
</main>
</body>
</html>

```

## Index.html

```

<!DOCTYPE html>
<html lang="en">
<head>
  <title>Login V20</title>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1"> <link
rel="icon" type="1/png" href="../static/images/icons/favicon.ico"/>
  <link rel="stylesheet" type="text/css" href="../static/vendor/bootstrap/css/bootstrap.min.css">
  <link rel="stylesheet" type="text/css" href="../static/fonts/font-awesome-4.7.0/css/font-awesome.min.css"> <link
rel="stylesheet" type="text/css" href="../static/fonts/Linearicons-Free-v1.0.0/icon-font.min.css"> <link rel="stylesheet"
type="text/css" href="../static/vendor/animate/animate.css">
  <link rel="stylesheet" type="text/css" href="../static/vendor/css-hamburgers/hamburgers.min.css"> <link
rel="stylesheet" type="text/css" href="../static/vendor/animation/css/animation.min.css"> <link
rel="stylesheet" type="text/css" href="../static/vendor/select2/select2.min.css">
  <link rel="stylesheet" type="text/css" href="../static/vendor/daterangepicker/daterangepicker.css"> <link
rel="stylesheet" type="text/css" href="../static/css/util.css"> <link rel="stylesheet" type="text/css"
href="../static/css/main.css">

</head>
<body>

  <div class="limiter">
    <div class="container-login100">
      <div class="wrap-login100 p-b-160 p-t-50">
        <form action="{{ url_for('abstract') }}" class="login100-form validate-form">
          <span class="login100-form-title p-b-43">
            Account Login
          </span>

```

```

<div class="wrap-input100 rs1 validate-input" data-validate = "Username is required">
  <input class="input100" type="text" name="username">
  <span class="label-input100">Username</span>
</div>

<div class="wrap-input100 rs2 validate-input" data-validate="Password is required">
  <input class="input100" type="password" name="pass">
  <span class="label-input100">Password</span>
</div>

<div class="container-login100-form-btn">
<button class="login100-form-btn">
  Sign in
</button>
</div>

<div class="text-center w-full p-t-23">
  <a href="#" class="txt1">
    Forgot password?
  </a>
</div>
</form>
</div>
</div>
</div>

<script src="../static/vendor/jquery/jquery-3.2.1.min.js"></script> <script
src="../static/vendor/animstition/js/animstition.min.js"></script> <script
src="../static/vendor/bootstrap/js/popper.js"></script> <script
src="../static/vendor/bootstrap/js/bootstrap.min.js"></script> <script
src="../static/vendor/select2/select2.min.js"></script>
<script src="../static/vendor/daterangepicker/moment.min.js"></script> <script
src="../static/vendor/daterangepicker/daterangepicker.js"></script> <script
src="../static/vendor/countdowntime/countdowntime.js"></script> <script
src="../static/js/main.js"></script>

</body>
</html>

```

# Upload.html

```
<!DOCTYPE html>
<html lang="en">

<head>

  <meta charset="utf-8">
  <meta name="viewport" content="width=device-width, initial-scale=1, shrink-to-fit=no">
  <meta name="description" content="">
  <meta name="author" content="">

  <title>Chronic Kidney Disease</title>

  <!-- Custom fonts for this theme -->
  <link href="../../static/vendor/fontawesome-free/css/all.min.css" rel="stylesheet" type="text/css">
  <link href="https://fonts.googleapis.com/css?family=Montserrat:400,700" rel="stylesheet" type="text/css"> <link
  href="https://fonts.googleapis.com/css?family=Lato:400,700,400italic,700italic" rel="stylesheet"
  type="text/css">

  <!-- Theme CSS -->
  <link href="../../static/css/freelancer.min.css" rel="stylesheet">

</head>

<body id="page-top">

  <!-- Navigation -->
  <nav class="navbar navbar-expand-lg bg-secondary text-uppercase fixed-top" id="mainNav">
    <div class="container">
      <a class="navbar-brand js-scroll-trigger" href="#page-top">Chronic Kidney Disease</a>

  <!-- About Section -->
  <section class="page-section bg-primary text-white mb-0"
  id="about"> <div class="container">
    <br>
    <br>
    <!-- About Section Heading -->
    <h2 class="text-center text-uppercase text-white">Upload Dataset</h2>

    <!-- Icon Divider -->
    <div class="divider-custom divider-light">
      <div class="divider-custom-line"></div>
      <div class="divider-custom-icon">
        <i class="fas fa-star"></i>
      </div>
      <div class="divider-custom-line"></div>
    </div>

    <!-- About Section Content -->
    <div class="row">
      <div class="col-lg-4 ml-auto" style="margin-right:250px;">
        <form action="http://localhost:5000/preview" name="fs" id="fs" method="post" enctype=multipart/form-
        data>
```

```

        <br/>
        <input type="file" name="datasetfile" id="file1" required
        /> <br/>
    <br/>
    <br/>
    <input type="submit" style="margin-right:250px" class=""
    value="Upload"> </form>
</div>

</section>

<!-- Copyright Section -->
<section class="copyright py-4 text-center text-white">
    <div class="container">

        </div>
    </section>

<!-- Scroll to Top Button (Only visible on small and extra-small screen sizes) -->
<div class="scroll-to-top d-lg-none position-fixed ">
    <a class="js-scroll-trigger d-block text-center text-white rounded" href="#page-top">
        <i class="fa fa-chevron-up"></i>
    </a>
</div>

<!-- Bootstrap core JavaScript -->
<script src="../static/vendor/jquery/jquery.min.js"></script>
<script src="../static/vendor/bootstrap/js/bootstrap.bundle.min.js"></script>

<!-- Plugin JavaScript -->
<script src="../static/vendor/jquery-easing/jquery.easing.min.js"></script>

<!-- Contact Form JavaScript -->
<script src="../static/js/jqBootstrapValidation.js"></script>
<script src="../static/js/contact_me.js"></script>

<!-- Custom scripts for this template -->
<script src="../static/js/freelancer.min.js"></script>

</body>

</html>

```

## Preview.html

```
<!DOCTYPE html>
<html lang="en">

<head>

  <meta charset="utf-8">
  <meta name="viewport" content="width=device-width, initial-scale=1, shrink-to-fit=no">
  <meta name="description" content="">
  <meta name="author" content="">

  <title>CHRONIC KIDNEY DISEASE</title>

  <!-- Custom fonts for this theme -->
  <link href="../static/vendor/fontawesome-free/css/all.min.css" rel="stylesheet" type="text/css">
  <link href="https://fonts.googleapis.com/css?family=Montserrat:400,700" rel="stylesheet" type="text/css"> <link
href="https://fonts.googleapis.com/css?family=Lato:400,700,400italic,700italic" rel="stylesheet"
type="text/css">

  <!-- Theme CSS -->
  <link href="../static/css/freelancer.min.css" rel="stylesheet">
  <style>
  #loading {
    background: url('../static/ajax-loader.gif') no-repeat center center;
    position: absolute;
    top: 0;
    left: 59;
    height: 100%;
    width: 86%;
    z-index: 9999999;
  }
  </style>
</head>
<body id="page-top">
  <!-- Navigation -->
  <nav class="navbar navbar-expand-lg bg-secondary text-uppercase fixed-top" id="mainNav">
    <div class="container">
      <a class="navbar-brand js-scroll-trigger" href="#page-top">Chronic Kidney
Disease</a> </ul>
    </div>
  </div>
</nav>
  <!-- Contact Section -->
  <section class="page-section" id="contact">
    <div class="container">
      <br>
      <br>
      <!-- Contact Section Heading -->
      <h2 class="text-center text-uppercase text-secondary mb-
0">Preview</h2> <!-- Icon Divider -->
      <div class="divider-custom">
        <div class="divider-custom-line"></div>
        <div class="divider-custom-icon">
          <i class="fas fa-star"></i>
        </div>
      </div>
    </div>
  </section>
</body>
</html>
```

```

        </div>
        <div class="divider-custom-line"></div>
    </div>
    <!-- Contact Section Form -->
    <div class="row" style="margin-left:-350px">
        <div class="col-lg-8 mx-auto">
            <!-- To configure the contact form email address, go to mail/contact_me.php and update the email address
in the PHP file on line 19. -->
            {{ df_view.to_html(classes="table striped",na_rep="-") |
            safe }} </div>
        </div>
    </section>
    <div class="form-group" style="padding:0px 250px 10px 40px;height:200px">
        <input style="margin-left:500px" type="button" onclick="hideLoader()" class="btn btn-primary"
value="Click to Train | Test" />
        <div id="loading" style="display:None;margin-
top:552950px"></div> </div>
    <section class="copyright py-4 text-center text-white">
        <div class="container">
            </div>
        </section>
    <!-- Scroll to Top Button (Only visible on small and extra-small screen sizes) -->
    <div class="scroll-to-top d-lg-none position-fixed ">
        <a class="js-scroll-trigger d-block text-center text-white rounded" href="#page-top">
            <i class="fa fa-chevron-up"></i>
        </a>
    </div>
    <script src="../static/vendor/jquery/jquery.min.js"></script>
    <script src="../static/vendor/bootstrap/js/bootstrap.bundle.min.js"></script> <script
src="../static/vendor/jquery-easing/jquery.easing.min.js"></script> <script
src="../static/js/jqBootstrapValidation.js"></script> <script
src="../static/js/contact_me.js"></script>
    <script src="../static/js/freelancer.min.js"></script>
    <script type='text/javascript' src='https://ajax.googleapis.com/ajax/libs/jquery/2.2.4/jquery.min.js'></script> <script
type='text/javascript'>
        function hideLoader() {
$( '#loading' ).show(0).delay(10000).hide(0,function(){
    alert("Training finished!");
    window.location = "{ url_for('home') }";
});
}
    </script>

</body>

</html>

```

## GITHUB LINK:

[https://github.com/CSE-DSU/TEAM\\_36\\_Chronic-Kidney-Disease-Prediction-Using-Machine-Learning](https://github.com/CSE-DSU/TEAM_36_Chronic-Kidney-Disease-Prediction-Using-Machine-Learning)

**PUBLISHED PAPER DETAILS**  
**PAPER ACCEPTANCE**

## FUNDING AND PUBLISHING PAPER DETAILS

<b>Paper Title</b>	Survey on Chronic Kidney Disease Prediction using Machine Learning
<b>Journal</b>	IEEE International conference on Advanced in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS 2023)
<b>Year of Publishing</b>	2023
<b>Abstract</b>	<p>Chronic kidney disease (CKD) is a significant global health problem with a high mortality and morbidity rate that also contributes to other ailments. Since there aren't any evident symptoms during the early stages of the condition, patients may ignore CKD. When the illness is discovered early, patients can receive quick treatment to stop the CKD from getting worse. Given their ability to accurately and quickly identify patterns, machine learning models may be able to help clinicians achieve this. The main objective of this study is to evaluate the performance of ML algorithms in comparison to other existing machine learning approaches in order to assess a person's propensity for CKD or not. The correlation of several methodologies included in this paper's performances can aid to launch a subsequent study beneficial for up-and-coming analysts in the sector.</p>
<b>Authors</b>	<p>Abhigna Srikara G  Akhila  Anjana KP  K Vaishali  Prof.Ranjani K</p>





K Vaishali &lt;vaishaliramesh2304@gmail.com&gt;

**International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS-2023) – Notification of paper submission status**

Microsoft CMT &lt;email@msr-cmt.org&gt;

Mon, Mar 13, 2023 at 5:10 PM

Reply-To: Dr Jalaja S &lt;jalajas@bit-bangalore.edu.in&gt;

To: K Vaishali &lt;vaishaliramesh2304@gmail.com&gt;

Dear Author(s),

Congratulations! The Program Committee of IEEE ICAECIS-2023 is pleased to inform you that your paper is accepted for presentation at IEEE ICAECIS-2023.

Your paper submission details are as below:

- Submission ID: 1155
- Title: Survey on Chronic Kidney Disease Prediction Using ML
- Status: Accept

We request you to carefully read through all of the terms and conditions in this email, and do the needful.

- Please ensure that you complete all the tasks mentioned below by the specified deadlines.
- Failure to do so will likely to forfeit your paper acceptance.
- Please continue to check the conference website <https://icaecis.com/> for all updates.

1. Only the designated corresponding author(s) for this paper will receive this letter and the enclosed material. If there are other authors, please share this correspondence with them. Names of all the authors will appear in the Conference Guide and the ICAECIS-2023 proceedings as will be given in the ICAECIS-2023 program database.

**NOTE:**

- It is mandatory for at least one of the authors to register and present the paper in the conference.
- The remaining authors can register for the conference either as an author or an attendee.
- Certificate will be issued only for the registered authors and registered attendees.
- Only the presented papers will be published in the IEEE ICAECIS-2023 conference proceedings and forwarded for publication in the on IEEE Xplore subsequently.

2. The review comments for your research paper can be viewed at the CMT paper submission page, <https://cmt3.research.microsoft.com/ICAECIS2023>.

- Please log in to view the reviews.
- Please incorporate the comments provided by the reviewers and revise your paper based on the comments.
- The camera-ready version of the paper must be prepared in accordance with the "Author Instructions" appended with this email.

3. The safety and well-being of all conference participants is our priority. As a precautionary measure it has been decided that IEEE ICAECIS-2023 will be held as hybrid mode (both online and offline) conference, from 19th to 21st April, 2023. Information regarding the virtual platform will be communicated in the coming days.

4. The conference program and additional information regarding the IEEE ICAECIS-2023 conference format will be posted on the conference website <https://icaecis.com/> at the earliest.

- We hereby request the authors to prepare presentation of their accepted paper (maximum of 12 slides) and a pre-recorded video of their presentation (maximum of 15 minutes).
- The PDF file of the presentation and the pre-recording of the paper presentation should be uploaded by 15th of April 2023 using the following link with the filename as your Paper ID.  
[https://drive.google.com/drive/folders/1Q6RwoFmoNetoCWBAZbgMTQ4s-QPA\\_qv47usp?usp=sharing](https://drive.google.com/drive/folders/1Q6RwoFmoNetoCWBAZbgMTQ4s-QPA_qv47usp?usp=sharing)

Final Paper Submission Instructions for Authors of ICAECIS-2023.

Authors, please complete the following steps by Saturday, March 25th, 2023.

1. Prepare your camera ready version, including the IEEE copyright notice
2. PDF eXpress Plus Online File Conversion/PDF Validation Tool
3. Complete and submit one IEEE copyright form per paper
4. Submit Camera ready Paper in both PDF and .doc file format
5. Register to the conference through the payment link provided

<https://mail.google.com/mail/u/0/?ik=6989d0609&view=pt&search=all&permmsgid=msg-f:1760252530717474397&simpl=msg-f:1760252530717...> 1/3



K Vaishali <vaishaliramesh2304@gmail.com>

## IEEE ICAECIS 2023 best paper and certificate information-

Microsoft CMT <email@msr-cmt.org>

Sat, Apr 22, 2023 at 12:05 PM

Reply-To: Dr Jalaja S <jalajas@bit-bangalore.edu.in>

To: K Vaishali <vaishaliramesh2304@gmail.com>

Dear K Vaishali,

Thank you for submitting your paper to IEEE ICAECIS 2023.

Your paper submission details referenced are per below:

Submission ID: 1155

Title: Survey on Chronic Kidney Disease Prediction Using ML

Status:Accept

Note:

1. Only Presented papers during the conference dated 19th, 20th and 21st April 2023 will be published in IEEE Xplore or Digital library.
2. Presented papers certificate will be sent by the end of first week of May 2023.
3. Certificate names will be retained for the presented paper certificate as per the CMT details. No further changes will be made.

### 4. ICAECIS - 2023 Best Paper awards

Paper details:

#### 1. 1st Prize (Paper ID: 1385)

"Hybrid RA\*2-Net: Residual Atrous Attention Network for Vessel Classification using Fundus Images  
Geetha Pavani P (Sankar Foundation Eye Hospital and Institute); Birendra Biswal (Gayatri Vidya Parishad College of Engineering(A))\*; Tapan Gandhi (Indian Institute of Technology Delhi); Krishna Talabhakthula (Sankar Foundation Eye Hospital )

#### 2. 2nd Prize (Paper ID: 1375)

DETECTION OF VIOLENT CONTENT IN VIDEOS USING AUDIO VISUAL FEATURES

Rishab K S (PES UNIVERSITY); Mayuravarsha P (PES UNIVERSITY)\*; Yashwal S Kanchan (PES UNIVERSITY ); Pranav M R (PES University); Roopa Ravish (PES University)

#### 2. 3rd Prize (Paper ID: 1328)

An innovative AI and image processing based ATM weapon, fraud detection and cybercrime prediction  
MANJUNATH K N (R V INSTITUTE OF TECHNOLOGY AND MANAGEMENT); ADARSHA H V SAGAR (RVITM)\*; Arshiya lubna (Presidency University)

We once again thank you that you considered IEEE ICAECIS 2023 as a forum to present your research.

Sincerely,

Dr. Jalaja S

Associate Dean, BIT-Bangalore,

IEEE ICAECIS 2023 TPC Chair.

Download the CMT app to access submissions and reviews on the move and receive notifications:

<https://apps.apple.com/us/app/conference-management-toolkit/id1532488001>

<https://play.google.com/store/apps/details?id=com.microsoft.research.cmt>

To stop receiving conference emails, you can check the 'Do not send me conference email' box from your User Profile.