

# Early Prediction Of Chronic Kidney Disease(CKD)

## Milestone 1: Project Initialization and Planning Phase

The "Project Initialization and Planning Phase" marks the project's outset, defining goals, scope, and predictions. This crucial phase establishes project parameters, identifies key team members, allocates resources, and outlines a realistic timeline. It also involves risk assessment and mitigation planning. Successful initiation sets the foundation for a well-organized and efficiently executed machine learning project, ensuring clarity, alignment, and proactive measures for potential challenges.

### Activity 1: Define Problem Statement

Problem Statement 1: A healthcare provider wants to identify individuals at risk of developing CKD. They can use machine learning models to analyze patient data, including age, gender, medical history, and lifestyle factors, to identify individuals with a high risk of developing CKD.

Problem Statement 2: A hospital wants to detect CKD in its early stages to provide timely treatment. They can use machine learning models to analyze patient data, such as blood pressure, blood glucose levels, and urine tests, to identify early signs of CKD and initiate treatment promptly.

Problem Statement 3: A healthcare provider wants to personalize treatment for patients with CKD based on their individual risk factors. They can use machine learning models to analyze patient data and predict the progression of CKD, helping them tailor treatment plans to each patient's needs.

Ref. template: [Click Here](#)

CKD Project Statement Report: [Click Here](#)

### Activity 2: Project Proposal (Proposed Solution)

The proposed project, "Early Prediction Of Chronic Kidney Disease" aims to leverage machine learning for more accurate CKD predictions. Using a comprehensive dataset including age, rbc, hemoglobin, anemia and diabetics, the project seeks to develop a predictive model optimizing CKD prediction.

Ref. template: [Click Here](#)

CKD Project Proposal Report: [Click Here](#)

### **Activity 3: Initial Project Planning**

Initial Project Planning involves outlining key objectives, defining scope, and identifying early prediction of CKD. It encompasses setting timelines, allocating resources, and determining the overall project strategy. During this phase, the team establishes a clear understanding of the dataset, formulates goals for analysis, and plans the workflow for data processing. Effective initial planning lays the foundation for a systematic and well-executed project, ensuring successful outcomes.

Ref. template: [Click Here](#)

CKD Project Planning Report: [Click Here](#)

## **Milestone 2: Data Collection and Preprocessing Phase**

The Data Collection and Preprocessing Phase involves executing a plan to gather relevant loan application data from Google drive, ensuring data quality through verification and addressing missing values. Preprocessing tasks include cleaning, encoding, and organizing the dataset for subsequent exploratory analysis and machine learning model development.

### **Activity 1: Data Collection Plan, Raw Data Sources Identified, Data Quality Report**

The dataset for " Early Prediction Of Chronic Kidney Disease " is sourced from Google drive. It includes patient details. Data quality is ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for predictive modeling.

Ref. template: [Click Here](#)

CKD Data Collection Report: [Click Here](#)

### **Activity 2: Data Quality Report**

The dataset for "Early Prediction Of Chronic Kidney Disease" is sourced from Google drive. It includes patient details. Data quality is ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for predictive modeling.

Ref. template: [Click Here](#)

CKD Data Quality Report: [Click Here](#)

### **Activity 3: Data Exploration and Preprocessing**

Data Exploration involves analyzing the Chronic Kidney Disease dataset to understand patterns, distributions, and outliers. Preprocessing includes handling missing values, scaling, and encoding categorical variables. These crucial steps enhance data quality, ensuring the reliability and effectiveness of subsequent analyses in the prediction of CKD.

Ref. template: [Click Here](#)

CKD Data Exploration and Preprocessing Report: [Click Here](#)

## **Milestone 3: Model Development Phase**

The Model Development Phase entails crafting a predictive model for loan approval. It encompasses strategic feature selection, evaluating and selecting models (Random Forest, Decision Tree, KNN, XGB), initiating training with code, and rigorously validating and assessing model performance for informed decision-making in the lending process.

### **Activity 1: Feature Selection Report**

The Feature Selection Report outlines the rationale behind choosing specific features (e.g., Age, RBC, Diabetics) for the CKD model. It evaluates relevance, importance, and impact on predictive accuracy, ensuring the inclusion of key factors influencing the model's ability to predict CKD.

Ref. template: [Click Here](#)

CKD Feature Selection Report: [Click Here](#) Activity

### **2: Model Selection Report**

The Model Selection Report details the rationale behind choosing Random Forest and Logistic Regression models for CKD prediction. It considers each model's strengths in handling complex relationships, interpretability, adaptability, and overall predictive performance, ensuring an informed choice aligned with project objectives.

Ref. template: [Click Here](#)

CKD Model Selection Report: [Click Here](#)

### **Activity 3: Initial Model Training Code, Model Validation and Evaluation Report**

The Initial Model Training Code employs selected algorithms on the Chronic Kidney Disease dataset, setting the foundation for predictive modeling. The subsequent Model Validation and Evaluation Report rigorously assesses model performance, employing metrics like accuracy and precision to ensure reliability and effectiveness in predicting CKD outcomes.

Ref. template: [Click Here](#)

CKD Model Development Phase Template: [Click Here](#)

## Milestone 4: Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

### Activity 1: Hyperparameter Tuning Documentation

The Random Forest model was selected for its superior performance, exhibiting high accuracy during hyperparameter tuning. Its ability to handle complex relationships, minimize overfitting, and optimize predictive accuracy aligns with project objectives, justifying its selection as the final model.

### Activity 2: Performance Metrics Comparison Report

The Performance Metrics Comparison Report contrasts the baseline and optimized metrics for various models, specifically highlighting the enhanced performance of the random Forest model. This assessment provides a clear understanding of the refined predictive capabilities achieved through hyperparameter tuning.

### Activity 3: Final Model Selection Justification

The Final Model Selection Justification articulates the rationale for choosing Random Forest as the ultimate model. Its exceptional accuracy, ability to handle complexity.

Ref. template: [Click Here](#)

CKD Model Optimization and Tuning Phase Report: [Click Here](#)

## Milestone 5: Project Files Submission and Documentation

For project file submission in Github, Kindly click the link and refer to the flow. [Click Here](#)

For the documentation, Kindly refer to the link. [Click Here](#)

## **Milestone 6: Project Demonstration**

In the upcoming module called Project Demonstration, individuals will be required to record a video by sharing their screens. They will need to explain their project and demonstrate its execution during the presentation.