Project Initialization and Planning Phase

Date	15 MARCH 2024
Team ID	LTVIP2024TMID25011
Project Title	Early Prediction Of Chronic Kidney Disease Using Machine Learning
Maximum Marks	3 Marks

Project Proposal (Proposed Solution) report

This project proposal outlines a solution to address Early Prediction Of Chronic Kidney Disease Using Machine Learning.Key features include a machine learning-based prediction model and real-time decision-making.

Project Overview	
Objective	This study proposes the following objectives: "To build and test a machine learning model for the early markers and risk factors of CKD with an intent to improve early diagnosing and early management strategies to increase patient's quality of living and decrease costs incurred in the healthcare setting.
Scope	This work proposes to design an ML system for classifying CKD employing medical data. Subprocesses cover data acquisition, data cleaning, and transformation, variable selection, model construction, and application to the healthcare domain, with emphasis on enhancing diagnostic performance and decision making aid for CKD.
Problem Statement	
Description	This project uses a machine learning approach to propose a method for the diagnosis of chronic kidney disease based on medical history data. In the context of CKD, it aims to improve diagnostic tools and care interventions for better decision-making process and overall patient management.
Impact	The model developed in this project for the diagnosis of chronic kidney diseases will transform the way the health sector prepares for and manages its patient load in the future by increasing detection rates and delivering better results from treatment. Its mission is lower the healthcare expenses and improve the patients' quality of life worldwide; this endeavor establishes a best practice for the utilization of AI in disease management.

Proposed Solution	
Approach	In the context of the CKD diagnosis project, specifications are for assembling clinical data on CKD patients, pre-processing this data (missing data imputation, normalization), analyzing the features' distribution, selecting appropriate machine learning algorithms, assessing their performance for clinical practice and deploying the solution.
Key Features	the integration of data from the patients' records, the results of the tests, advanced data preprocessing (handling missing values in the dataset, normalization of data), the selection of the critical features of the dataset such as creatinine level or demographic characteristics, and the model selection and tuning such as Decision Trees and Support Vector Machine (SVM). The solution's specific goal is to provide high diagnostic accuracy

Resource Requirements

Resource Type	Description	Specification/Allocation
Hardware		
Computing Resources	CPU/GPU specifications, number of cores	T4 GPUs
Memory	RAM specifications	8 GB
Storage	Disk space for data, models, and logs	1 TB SSD
Software		
Frameworks	Python frameworks	Flask
Libraries	Additional libraries	scikit-learn, pandas,pickle numpy,collections ,sklearn matplotlib,seaborn,missingno

Development Environment	IDE, version control	Jupyter Notebook, Git,vscode
Data		
Data	Source, size, format	Kaggle dataset, 10,000 images,CSV,health department,patient datasets