

Project Initialization and Planning Phase

Date	5 July 2024
Team ID	SWTID1720082525
Project Title	Early Prediction of Chronic Kidney Disease Using Machine Learning
Maximum Marks	3 Marks

Project Proposal (Proposed Solution) template

This project aims to develop a machine learning model for early prediction of chronic kidney disease using adult patient data from electronic health records. By analysing past medical history, lab results, and demographic information, the model will provide real-time risk assessments, enabling early intervention. Key features include high accuracy, EHR integration, and interpretable results for clinicians. The goal is to improve patient outcomes and reduce healthcare costs associated with late-stage CKD diagnosis and treatment.

Project Overview	
Objective	To develop a machine learning model that can accurately predict the onset of chronic kidney disease (CKD) in its early stages, enabling timely intervention and improved patient outcomes.
Scope	The project scope encompasses developing a predictive model for chronic kidney disease using EHR data from adults aged 18 and above, collected over the past 5 years. The model will be implemented and tested in a clinical setting to validate its effectiveness in real-world healthcare environments.
Problem Statement	
Description	Current CKD detection methods often lead to late diagnoses, negatively impacting patient outcomes and healthcare efficiency. This project leverages machine learning to analyze patient data, enabling earlier and more accurate identification of at-risk individuals. The solution aims to improve diagnostic accuracy, facilitate timely interventions, and enhance overall patient care quality in CKD management.
Impact	Early prediction of CKD can:

	<ul style="list-style-type: none"> • Improve patient outcomes through timely intervention • Reduce healthcare costs associated with late-stage treatments • Enhance quality of life for at-risk individuals • Decrease the burden on healthcare systems
Proposed Solution	
Approach	The approach involves collecting anonymized patient data from EHRs, preprocessing it, and selecting relevant features for CKD prediction. Multiple machine learning algorithms will be implemented and compared to develop the best performing model, which will then be validated using cross-validation techniques. A user-friendly interface will be created for clinical integration, allowing healthcare providers to input data and receive risk assessments. Finally, the model will undergo pilot testing in a clinical setting to gather professional feedback and refine its performance.
Key Features	<ol style="list-style-type: none"> 1. Machine learning-based CKD risk assessment model 2. Real-time high-risk patient identification 3. Adaptive learning for evolving medical knowledge 4. EHR integration for comprehensive data analysis 5. User-friendly interface for healthcare providers

Resource Requirements

Resource Type	Description	Specification/Allocation
Hardware		
Computing Resources	CPU/GPU specifications, number of cores	e.g., 2 x NVIDIA V100 GPUs
Memory	RAM specifications	e.g., 8 GB
Storage	Disk space for data, models, and logs	e.g., 1 TB SSD
Software		
Frameworks	Python frameworks	e.g., Flask
Libraries	Additional libraries	e.g., scikit-learn, pandas, numpy
Development Environment	IDE, version control	e.g., Jupyter Notebook, Git

Data		
Data	Source, size, format	Kaggle/Skill Wallet dataset,10,000images