Project Initialization and Planning Phase

Date	14 June 2025
Team ID	SWTID1749713922
Project Title	Early prediction for chronic kidney disease detection: A progressive approach to health management
Maximum Marks	3 Marks

Project Proposal

This project proposal outlines a solution to address a specific problem. With a clear objective, defined scope, and a concise problem statement, the proposed solution details the approach, key features, and resource requirements, including hardware, software, and personnel.

Project Overview	
Objective	To develop a machine learning model that accurately predicts early-stage kidney disease using patient medical data.
Scope	The project focuses on supervised learning models trained on structured patient data to predict the presence of kidney disease, and includes data preprocessing, model evaluation, and feature importance analysis. It excludes deployment and real-time integration.
Problem Statement	
Description	Chronic Kidney Disease (CKD) is often diagnosed late, leading to severe health complications. Early detection is critical but often missed due to a lack of predictive tools in primary diagnosis stages.
Impact	A reliable ML-based early warning system can assist healthcare providers in identifying at-risk patients earlier, improving patient outcomes and reducing long-term treatment costs.
Proposed Solution	
Approach	The solution involves training multiple machine learning classifiers on a medical dataset. Then feature selection techniques are applied to improve model performance. The best-performing model will be selected based evaluation

	metrics.
Key Features	Data preprocessing Exploratory data analysis Data encoding Model training Application building

Resource Requirements

Resource Type	Description	Specification/Allocation
Hardware		
Computing Resources	CPU/GPU specifications, number of cores	Intel Core i5
Memory	RAM specifications	12 GB
Storage	Disk space for datasets and models	512 GB SSD
Software		
Frameworks	ML and web frameworks	Python, Flask
Libraries	Libraries for ML and data manipulation	scikit-learn, pandas, NumPy, seaborn, matplotlib
Development Environment	Tools for writing and testing code	Jupyter Notebook, GitHub
Data		
Data	Structured health records	CKD dataset, 400 samples (CSV format)