Disease Prediction Using Machine Learning

Milestone 1: Project Initialization and Planning Phase

The "Project Initialization and Planning Phase" marks the project's outset, defining goals, scope, and stakeholders. 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: The early prediction of diseases can significantly improve patient outcomes, reduce healthcare costs, and enhance the quality of life. With the advent of big data, machine learning, and advanced analytics, there is a growing potential to predict diseases before they manifest clinically. This predictive capability can aid in proactive treatment, preventive measures, and personalized healthcare. In the realm of healthcare, accurately predicting the onset and progression of diseases is crucial for early intervention and improved patient outcomes. Despite advancements in medical technology and data analytics, current methods often lack precision and timeliness, leading to delayed diagnoses and suboptimal treatment plans. This project aims to develop a robust and scalable disease prediction model that leverages patient data, including medical history, genetic information, lifestyle factors, and real-time health indicators, to forecast the likelihood of disease occurrence with high accuracy. The goal is to enhance predictive capabilities, reduce healthcare costs, and ultimately improve patient care through timely and personalized interventions.

Disease Prediction Problem Statement Report: [Click Here](#)

Activity 2: Project Proposal (Proposed Solution)

This project proposal outlines a solution to address the problem of early disease detection through machine learning. With a clear objective to develop a predictive model for assessing disease risk based on symptoms, lifestyle factors, and health data, the proposal defines the scope of the project, including data collection, model development, and deployment. The proposed solution details the approach to be used, key features of the

model, and specifies the resource requirements including hardware, software, and personnel. By creating an accurate and user-friendly tool, the project aims to enable proactive health management and improve early disease detection.

Disease Prediction Project Proposal Report: Click Here

Activity 3: Initial Project Planning
Initial Project Planning involves outlining key objectives, defining scope, and identifying stakeholders for a loan approval system. 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.

Disease Prediction 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 Kaggle, 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 "Disease Prediction Using Machine Learning" is sourced from Kaggle. It includes data quality to Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Disease Prediction Data Collection Report: Click Here

Activity 2: Data Quality Report
The dataset for "Disease Prediction Using Machine Learning" is sourced from Kaggle. It includes data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

Disease Prediction Data Quality Report: Click Here

Activity 3: Data Exploration and Preprocessing
Dataset variables will be statistically analyzed to identify patterns and outliers, with Python employed for preprocessing tasks like normalization and feature engineering. Data cleaning will address missing values and outliers, ensuring quality for subsequent analysis and modeling, and forming a strong foundation for insights and predictions

Disease Prediction 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, SVC), 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 forthcoming update, each feature will be accompanied by a brief description. Users will indicate whether it's selected or not, providing reasoning for their decision. This process will streamline decision-making and enhance transparency in feature selection.

Disease Prediction Feature Selection Report: [Click Here](#)

Activity 2: Initial Model Training Code, Model Validation and Evaluation Report
The Initial Model Training Code employs selected algorithms on the Disease Prediction 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 Diseases

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

Activity 3: Model Selection Report
The Model Selection Report details the rationale behind choosing Random Forest, Decision Tree, KNN, and SVC models for loan approval 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.

Disease Prediction Model Selection Report: [Click Here](#)

Milestone 3: 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 .clickhere

Activity 1: Hyperparameter Tuning Documentation

The knn 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 knn 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 knn as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successful hyperparameter tuning align with project objectives, ensuring optimal loan approval predictions.

Disease Prediction 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.