

Revolutionizing Liver Care – Predicting Liver Cirrhosis using advanced 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 goal of this project is to develop a predictive model that accurately identifies patients at high risk of developing cirrhosis of the liver. We'll use machine learning techniques to analyze patient data and predict the likelihood of cirrhosis based on relevant features.

Ref. template: [Click Here](#)

SmartLender Problem Statement Report: [Click Here](#)

Activity 2: Project Proposal (Proposed Solution)

Liver cirrhosis is a chronic and progressive liver disease that affects millions of people worldwide. It is a leading cause of morbidity and mortality, with a significant impact on healthcare systems and society. Early detection and prediction of cirrhosis are crucial for effective management and prevention of complications. However, current diagnostic methods are often unreliable and inaccurate

Ref. template: [Click Here](#)

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

Activity 3: Initial Project Planning

This initial planning provides a structured approach to developing a liver cirrhosis prediction model using advanced machine learning techniques. It covers all the essential steps from data collection to deployment, ensuring a comprehensive and practical solution.

Ref. template: [Click Here](#)

SmartLender 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 cirrhosis

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

This comprehensive data collection plan, identification of raw data sources, and detailed data quality report ensure that the data used for developing the liver cirrhosis prediction model is reliable, accurate, and suitable for advanced machine learning analysis.

40

Ref. template: [Click Here](#)

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

Activity 2: Data Quality Report

The dataset for Liver Cirrhosis This data quality report provides a comprehensive assessment of the quality of the data collected for the liver cirrhosis prediction model. The report covers various dimensions of data quality, including completeness, consistency, accuracy, timeliness, validity, and reliability.

Ref. template: [Click Here](#)

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

Activity 3: Data Exploration and Preprocessing

Data Exploration involves

- **Objective:** Understand the structure, attributes, and basic statistics of the dataset.
- **Actions:**
 - Load the dataset into a data analysis environment (e.g., Python using pandas).
 - Display the first few rows of the dataset to inspect the data format and columns.
 - Use `info()` method to get a summary of the dataset including data types and non-null counts

Ref. template: [Click Here](#)

SmartLender 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 Predicting process.

Activity 1: Feature Selection Report

This report details the process and results of feature selection for the liver cirrhosis prediction model. The objective is to identify the most relevant features that significantly contribute to predicting liver cirrhosis. Effective feature selection can improve model performance, reduce overfitting, and enhance interpretability.

Ref. template: [Click Here](#)

SmartLender Feature Selection Report: [Click Here](#)

Activity 2: Model Selection Report

The Model Selection Report This report outlines the process of selecting the most suitable machine learning model for predicting liver cirrhosis. Several models were evaluated based on their performance metrics, and the best-performing model was identified for further tuning and deployment.

Ref. template: [Click Here](#)

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

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

1. **Accuracy:**
 - The proportion of correctly predicted instances out of the total instances.
2. **Precision:**
 - The proportion of true positive predictions out of the total predicted positives.
3. **Recall:**
 - The proportion of true positive predictions out of the total actual positives.
4. **F1-Score:**
 - The harmonic mean of precision and recall.
5. **ROC-AUC:**
 - The area under the Receiver Operating Characteristic curve, indicating the model's ability to distinguish between classes.

The confusion matrix is also provided to visualize the number of true positive, true negative, false positive, and false negative predictions.

Ref. template: [Click Here](#)

SmartLender 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 Gradient Hyperparameter tuning aims to optimize the performance of the Random Forest Classifier by systematically adjusting its parameters. This process helps in achieving better accuracy, precision, recall, and other evaluation metrics by finding the optimal combination of hyperparameters.

Activity 2: Performance Metrics Comparison Report:

The Final This report compares the performance metrics of the Random Forest Classifier before and after hyperparameter tuning. The goal is to demonstrate the improvement in the model's performance due to tuning and to highlight the key metrics used for comparison.

Activity 3: Final Model Selection Justification

This document provides a detailed justification for selecting the final model for predicting liver cirrhosis. The selection process involved evaluating multiple models, tuning hyperparameters, and comparing performance metrics to ensure the chosen model provides the best predictive performance and generalizability.

Ref. template: [Click Here](#)

SmartLender 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

The project aims to predict liver cirrhosis using advanced machine learning techniques. This demonstration will showcase the steps involved in data preprocessing, model training (specifically with the Random Forest Classifier), hyperparameter tuning, and evaluating the final model's performance metrics.