

# Polycystic Ovary Syndrome

## 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: Predicting the Polycystic Ovary Syndrome accurately based on particular symptoms is essential for early treatment and medical care of the patients and prevention of the disease. The goal is to develop a predictive model that can predict the disease using available dataset features, contributing to better medical improvements and discovery of new treatment techniques. Ref. template: [Click Here](#)

Polycystic Ovary syndrome Classification Problem Statement Report: [Click Here](#)

### Activity 2: Project Proposal (Proposed Solution)

The proposed project, "Polycystic Ovary Syndrome (PCOS) Classification Using Machine Learning," aims to leverage machine learning for accurate classification of PCOS. Using a comprehensive dataset that includes features such as age, Weight Gain, RBS, luteinizing hormone (LH), follicle-stimulating hormone (FSH), testosterone levels, and other relevant clinical and hormonal data, the project seeks to develop a predictive model to identify and classify PCOS in patients. This model will assist healthcare professionals in early diagnosis and personalized treatment planning, thereby improving patient outcomes and contributing to better understanding and management of PCOS.

Ref. template: [Click Here](#)

Polycystic Ovary syndrome Classification Project Proposal Report: [Click Here](#)

### Activity 3: Initial Project Planning

Initial Project Planning involves outlining key objectives, defining scope, and identifying stakeholders for the abalone age prediction project. 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.

Ref. template: [Click Here](#)

Polycystic Ovary syndrome Classification Initial 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 PCOS 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 "Polycystic Ovary Syndrome (PCOS) classification Using Machine Learning" will be sourced from public medical repositories and clinical studies. It includes clinical and hormonal data such as age, Weight Gain, RBS, luteinizing hormone (LH), follicle-stimulating hormone (FSH), testosterone levels, and other relevant patient information. Data quality will be ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for developing a robust predictive model for PCOS classification.

Ref. template: [Click Here](#)

Polycystic Ovary Syndrome Classification Data Collection Report: [Click Here](#)

### Activity 2: Data Quality Report

The dataset for "Polycystic Ovary Syndrome (PCOS) classification Using Machine Learning" will be sourced from public medical repositories and clinical studies. It includes clinical and hormonal data such as age, RBS, Weight Gain, luteinizing hormone (LH), follicle-stimulating hormone (FSH), and other relevant patient information. Data quality will be ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for developing a robust predictive model for PCOS classification.

Ref. template: [Click Here](#)

Abalone Age Prediction Data Quality Report: [Click Here](#)

### Activity 3: Data Exploration and Preprocessing

Data exploration involves analyzing the PCOS dataset to understand patterns, distributions, and outliers. Preprocessing includes handling missing values, scaling, and encoding categorical variables. These crucial steps enhance the dataset's quality, ensuring it is ready for model training

and testing phases.

Ref. template: [Click Here](#)

### **Polycystic Ovary Syndrome Classification 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, Linear Regression), initiating training with code, and rigorously validating and assessing model performance for predicting Abalone's Age accurately.

### **Activity 1: Feature Selection Report**

The Feature Selection Report outlines the rationale behind choosing specific features (e.g., age, RBS, Weight gain, luteinizing hormone (LH), follicle-stimulating hormone (FSH), BP\_Diastolic) for the PCOS classification model. It evaluates relevance, importance, and impact on predictive accuracy, ensuring the inclusion of key factors influencing the model's ability to classify PCOS accurately.

Ref. template: [Click Here.](#)

**Polycystic Ovary Syndrome Classification Feature Selection Report:** [Click Here](#)

### **Activity 2: Model Selection Report**

The Model Selection Report details the rationale behind choosing Random Forest, Decision Tree, Linear Regression models for PCOS Classification. It considers each model's strengths in handling complex relationships, interpretability, adaptability, and overall predictive performance, ensuring an informed choice aligned with project objectives.

Ref. template: [Click Here](#)

**Polycystic Ovary Syndrome Model Selection Report:** [Click Here](#)

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

The Initial Model Training Code employs selected algorithms on the loan approval dataset, setting the foundation for predictive modelling. The subsequent Model Validation and Evaluation Report rigorously assesses model performance, employing metrics like accuracy and precision to ensure reliability and effectiveness in classifying PCOS.

Ref. template: [Click Here](#)

**Polycystic Ovary Syndrome Model Development Phase Report:** [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 Decision Tree model 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 Random Forest Classifier 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 Random Forest Classifier as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successful hyperparameter tuning align with project objectives, ensuring optimal Polycystic Ovary Syndrome Classification.

**Ref. template:** [Click Here.](#)

**PCOS 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.