PCOS Care
YOUR PARTNER FOR WELLNESS

PCOS Care: A Web Application for Diagnosing PCOS Patients Using Machine Learning

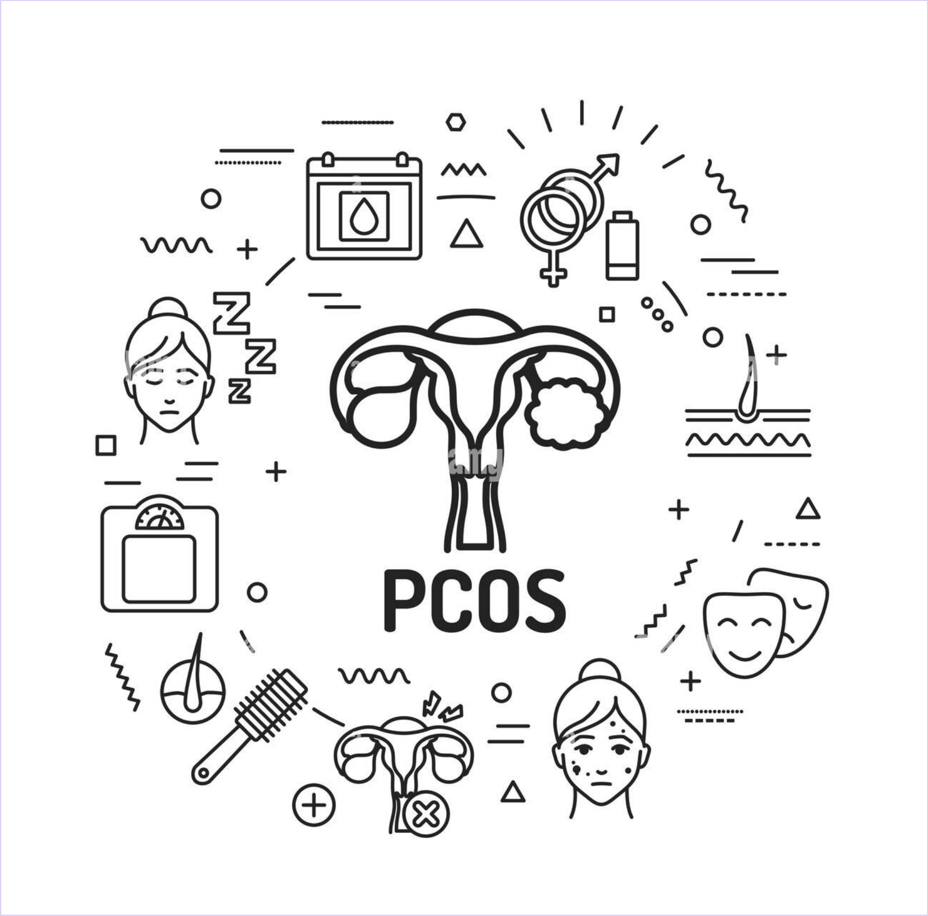
Interim Progression Demonstration (IPD)

Student: Rachel Naadia Cooray (19564441)
Supervisor: Dr Malarvizhi Kaniappan Chinnathai
Degree: BSc (Hons) Computer Science

13th February 2025

Content

- Recap of the Problem
- Project Stakeholders - Roles & Influence
- Requirements & their Alignment with aims
- Formal Requirements Documentation
- System Architecture & Design
- Development Progress
- Key Features & Functionalities Developed
- Challenges Faced & Solutions
- Standout Achievements
- Prototype Overview
- Upcoming Milestones
- References



Recap of the Problem

"Years passed and I was going to different GPs and no one could get to the bottom of the issue."
(Battiste, BBC, 2024).

OVERVIEW

This project develops a web-based application that uses machine learning (ML) to assess Polycystic ovary syndrome (PCOS) risk and provide educational resources.

SIGNIFICANCE

Around 70% of women with PCOS remain undiagnosed due to diagnostic challenges and limited healthcare access (Adla et al., 2021).

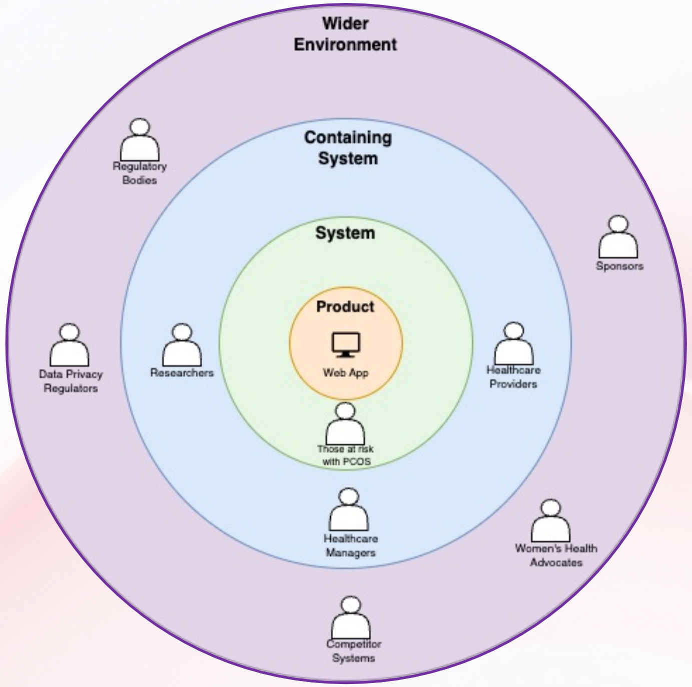
Early detection through this tool can improve health outcomes and empower users to seek timely medical advice.

PROBLEM & SOLUTION

PCOS diagnosis is complex, and time consuming. This project addresses the gap by creating a user-friendly application for preliminary risk assessment, supporting awareness and early intervention.

Project Stakeholders

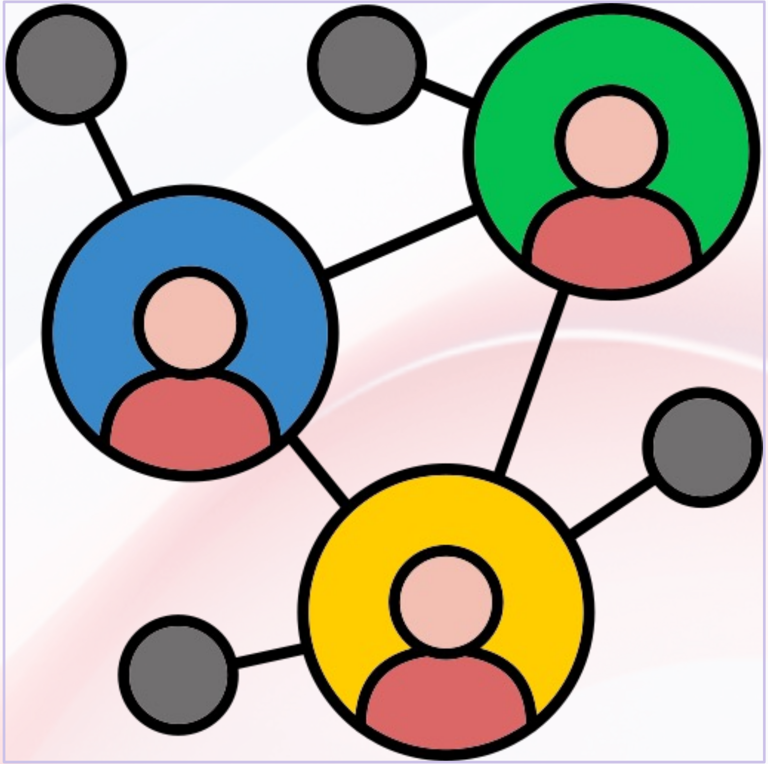
Stakeholder	Description
Users	Women at risk of or experiencing PCOS symptoms.
Healthcare Providers	Endocrinologists and gynaecologists diagnosing PCOS.
Healthcare Managers	Those in charge of medical facilities and patient care.
Researchers	Those studying PCOS in academia and industry.
Regulatory Bodies	Organizations that establish rules and regulations for healthcare applications.
Data Privacy Regulators	Authorities responsible for ensuring privacy laws are followed.
Women's Health Advocates	Groups working to raise awareness about women's health issues.
Competitor Systems	Other platforms or studies offering PCOS management tools.
Sponsors	Potential investors supporting the application's development.



Onion Model (Self composed)

Roles & Influence

Stakeholder	Role and Influence
Users	Provide input data and receive predictions.
Healthcare Providers	Use app for preliminary assessments.
Healthcare Managers	Oversee implementation in medical settings.
Researchers	Analyse data to enhance prediction models.
Regulatory Bodies	Ensure compliance with healthcare laws.
Data Privacy Regulators	Enforce data protection and privacy policies.
Women's Health Advocates	Enhance awareness and outreach.
Competitor Systems	Set industry benchmarks and drive innovation.
Sponsors	Provide funding and resources for development.



Requirements & their Alignment with aims

Requirement	Requirement Type	Supporting Objective
Data Input Form	Functional	Enable users to input health data and symptoms
Prediction Model	Functional	Provide accurate PCOS predictions
Results Display	Functional	Display predictions in a user-friendly manner
Information Page	Functional	Provide educational resources on PCOS
Performance	Non-functional	Efficient and responsive system
Usability	Non-functional	Application is easy to use
Security	Non-functional	Safeguard personal data entered by users
Compatibility	Non-functional	Application runs on multiple devices

Functional Requirements

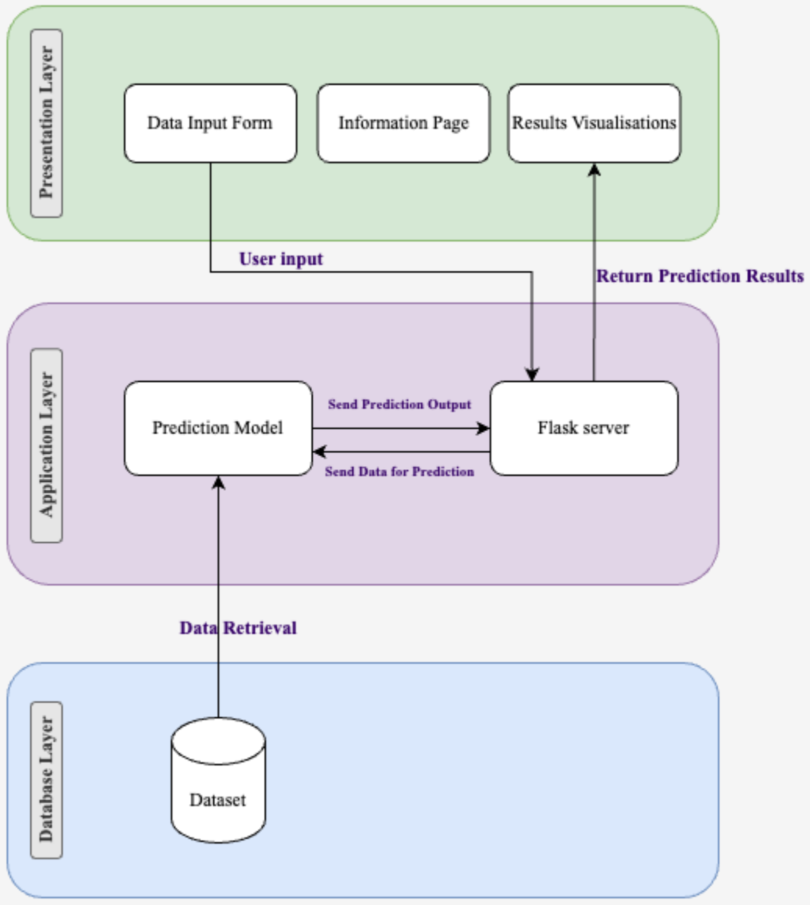
Functional Requirement	Description	Progress
Data Input Form	Users can input symptoms, medical data	Implemented
Prediction Model	Use ML to predict likelihood of PCOS	Implemented
Results Display	Graphical display of prediction results	Statistical testing implemented
Information Page	Provide information on PCOS symptoms, causes, and management	Basic UI implemented

Non-Functional Requirements

Non-Functional Requirement	Description	Progress
Performance	Ensure fast response times during input, prediction, and display	Pages load fast Further testing needed when pages are linked
Usability	Provide an intuitive, user-friendly interface	Basic frontend implemented Pending usability testing to get user feedback
Security	User data is not stored or shared with any third parties	User data not stored or shared
Compatibility	The application runs seamlessly on multiple devices and browsers	Tested on Safari and Chrome

System Architecture & Design

System Architecture Diagram



System Architecture Diagram (Self composed)

Key Reasons for a 3-Tier System

Scalability

Each layer can be modified independently for future improvements.

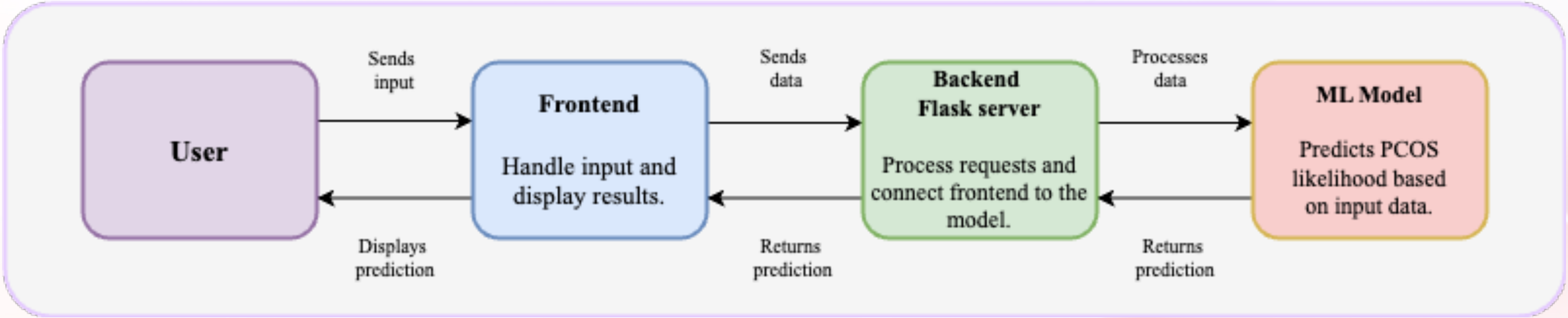
Maintainability

Easier debugging and updating of individual components.

Performance

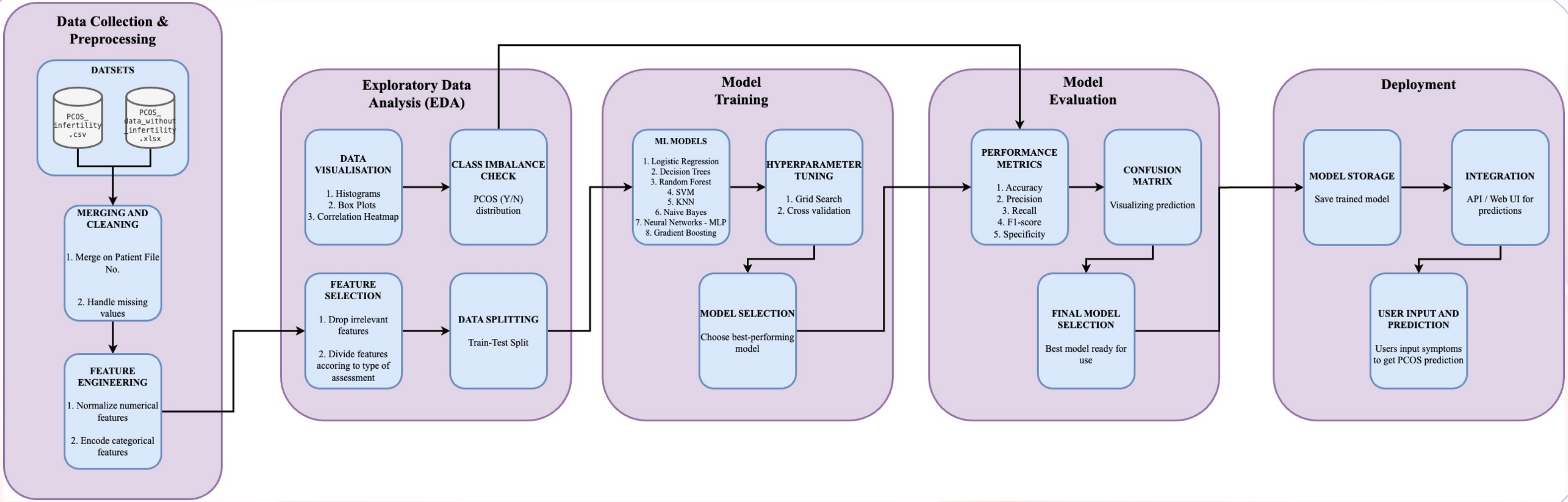
Efficient handling of user input, model processing, and data storage.

High level diagram of the application



High Level Diagram (Self composed)

High level diagram of the model



Data Collection & Preprocessing

DATSETS

PCOS_ infertility .csv

PCOS_ data_without infertility .xlsx

MERGING AND CLEANING

1. Merge on Patient File No.

2. Handle missing values

FEATURE ENGINEERING

1. Normalize numerical features

2. Encode categorical features

Exploratory Data Analysis (EDA)

DATA VISUALISATION

1. Histograms
2. Box Plots
3. Correlation Heatmap

CLASS IMBALANCE CHECK

PCOS (Y/N) distribution

FEATURE SELECTION

1. Drop irrelevant features

2. Divide features accoring to type of assessment

DATA SPLITTING

Train-Test Split

Model Training

ML MODELS

1. Logistic Regression
2. Decision Trees
3. Random Forest
4. SVM
5. KNN
6. Naive Bayes
7. Neural Networks - MLP
8. Gradient Boosting

HYPERPARAMETER TUNING

1. Grid Search
2. Cross validation

MODEL SELECTION

Choose best-performing model

Model Evaluation

PERFORMANCE METRICS

1. Accuracy
2. Precision
3. Recall
4. F1-score
5. Specificity

CONFUSION MATRIX

Visualizing prediction

FINAL MODEL SELECTION

Best model ready for use

Deployment

MODEL STORAGE

Save trained model

INTEGRATION

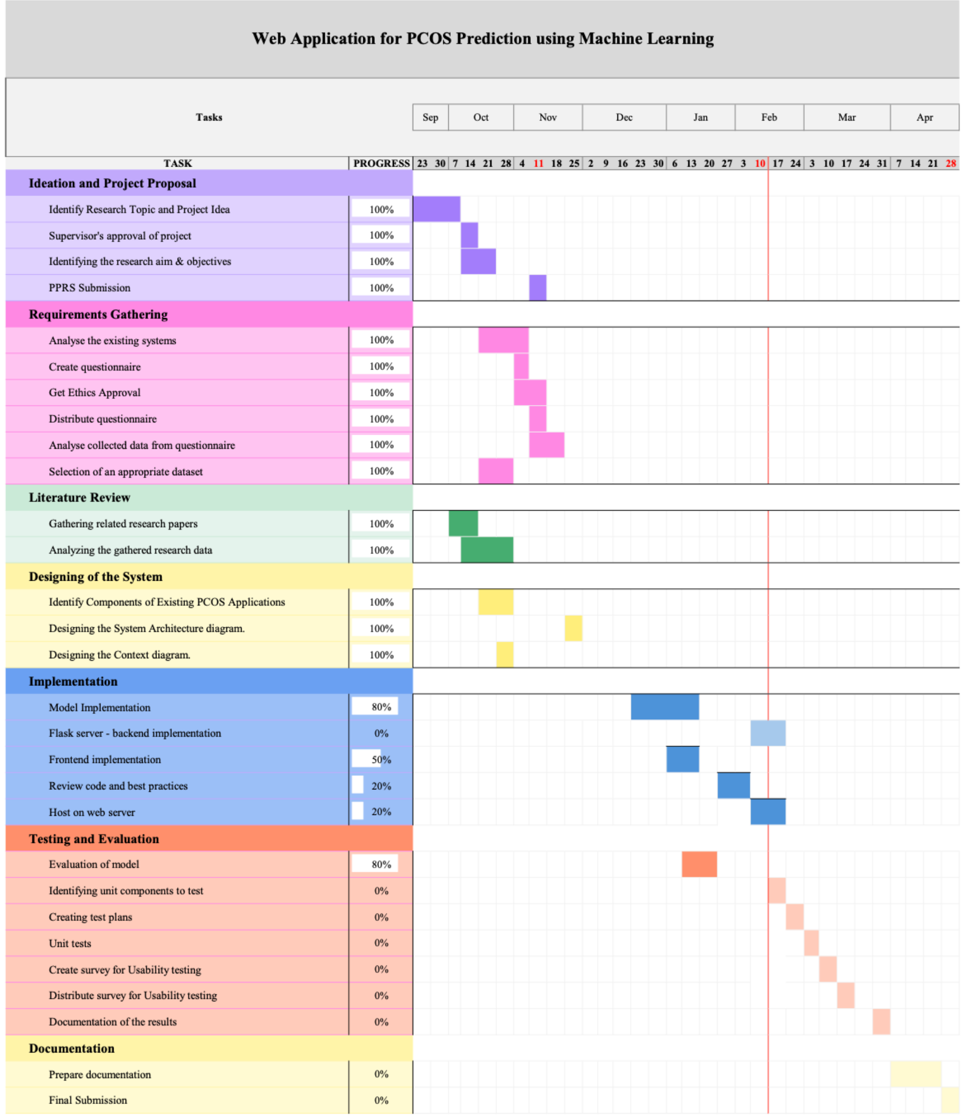
API / Web UI for predictions

USER INPUT AND PREDICTION

Users input symptoms to get PCOS prediction

High Level Diagram for model (Self composed)

Development Progress



Web Application for PCOS Prediction using Machine Learning

TASK	PROGRESS
Ideation and Project Proposal	
Identify Research Topic and Project Idea	100%
Supervisor's approval of project	100%
Identifying the research aim & objectives	100%
PPRS Submission	100%
Requirements Gathering	
Analyse the existing systems	100%
Create questionnaire	100%
Get Ethics Approval	100%
Distribute questionnaire	100%
Analyse collected data from questionnaire	100%
Selection of an appropriate dataset	100%
Literature Review	
Gathering related research papers	100%
Analyzing the gathered research data	100%
Designing of the System	
Identify Components of Existing PCOS Applications	100%
Designing the System Architecture diagram.	100%
Designing the Context diagram.	100%
Implementation	
Model Implementation	80%
Flask server - backend implementation	0%
Frontend implementation	50%
Review code and best practices	20%
Host on web server	20%
Testing and Evaluation	
Evaluation of model	80%
Identifying unit components to test	0%
Creating test plans	0%
Unit tests	0%
Create survey for Usability testing	0%
Distribute survey for Usability testing	0%
Documentation of the results	0%
Documentation	
Prepare documentation	0%
Final Submission	0%

Progress

- Basic UI designed.
- Model training completed.

Additions

- To implement web hosting.

Deviations

- Additional effort in January realigned the project with the schedule.

Link to Gantt Chart

Key Features & Functionalities Developed

01

Prediction Model

02

UI Design Mock-ups

03

Initial Frontend Pages

Link to Prototype Files

01. Prediction Model

Data Preprocessing

- Merged datasets, handled missing values, and cleaned dataset.

Feature Selection

- Used Correlation, ANOVA and Chi-Square tests for relevant features.

Model Selection

- Trained Classification models such as Logistic Regression, SVM, Random Forest, KNN and Neural Networks.

Train-Test Split

- Models trained and tested on these data.

Sensitivity Analysis

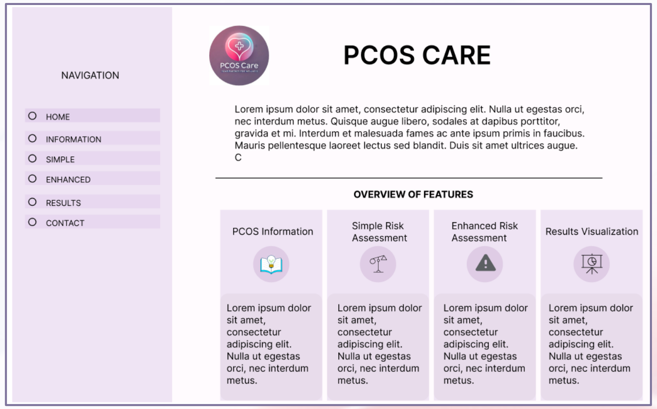
- Tested different feature sets and data splits of 80:20, 85:15, and 90:10 for optimal accuracy.

Performance Metrics

- Assessed using Accuracy, Precision, Recall, Specificity and F1 Score.

02. UI Design Mock-ups

- Designed using Figma, following the format of Streamlit.
- Includes wireframes and mock-ups.
- Focused on clarity, accessibility, and ease of use.
- **Data Input Form** for seamless entry.
- **Prediction Results Page** with interactive graphs.
- **Information Resources Page** for guidance.



PCOS CARE

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla ut egestas orci, nec interdum metus. Quisque augue libero, sodales at dapibus porttitor, gravida et mi. Interdum et malesuada fames ac ante ipsum primis in faucibus. Mauris pellentesque laoreet lectus sed blandit. Duis sit amet ultrices augue. C

OVERVIEW OF FEATURES

PCOS Information

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla ut egestas orci, nec interdum metus.

Simple Risk Assessment

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla ut egestas orci, nec interdum metus.

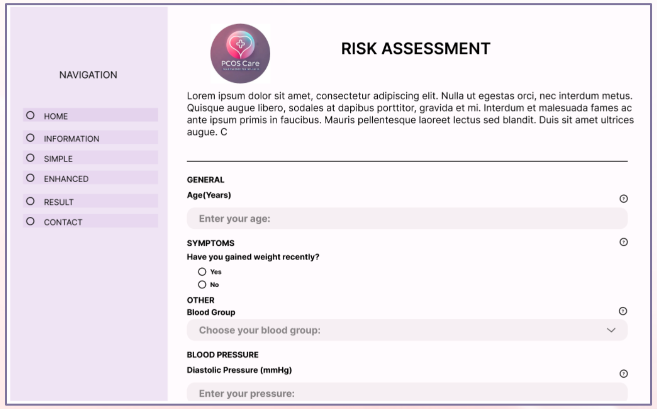
Enhanced Risk Assessment

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla ut egestas orci, nec interdum metus.

Results Visualization

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla ut egestas orci, nec interdum metus.

Mock-up 1 (Self composed)



RISK ASSESSMENT

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla ut egestas orci, nec interdum metus. Quisque augue libero, sodales at dapibus porttitor, gravida et mi. Interdum et malesuada fames ac ante ipsum primis in faucibus. Mauris pellentesque laoreet lectus sed blandit. Duis sit amet ultrices augue. C

GENERAL

Age(Years)

Enter your age:

SYMPTOMS

Have you gained weight recently?

○ Yes
○ No

OTHER

Blood Group

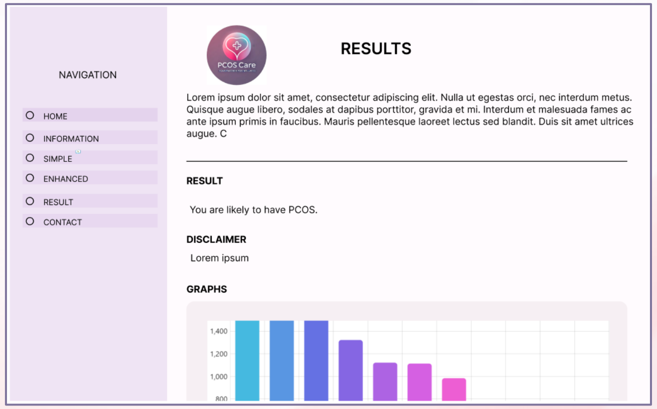
Choose your blood group:

BLOOD PRESSURE

Diastolic Pressure (mmHg)

Enter your pressure:

Mock-up 2 (Self composed)



RESULTS

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla ut egestas orci, nec interdum metus. Quisque augue libero, sodales at dapibus porttitor, gravida et mi. Interdum et malesuada fames ac ante ipsum primis in faucibus. Mauris pellentesque laoreet lectus sed blandit. Duis sit amet ultrices augue. C

RESULT

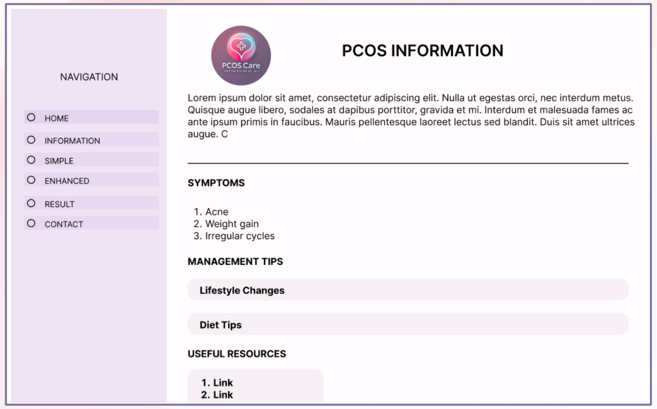
You are likely to have PCOS.

DISCLAIMER

Lorem ipsum

GRAPHS

Mock-up 3 (Self composed)



PCOS INFORMATION

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla ut egestas orci, nec interdum metus. Quisque augue libero, sodales at dapibus porttitor, gravida et mi. Interdum et malesuada fames ac ante ipsum primis in faucibus. Mauris pellentesque laoreet lectus sed blandit. Duis sit amet ultrices augue. C

SYMPTOMS

1. Acne
2. Weight gain
3. Irregular cycles

MANAGEMENT TIPS

Lifestyle Changes

Diet Tips

USEFUL RESOURCES

1. Link
2. Link

Mock-up 4 (Self composed)

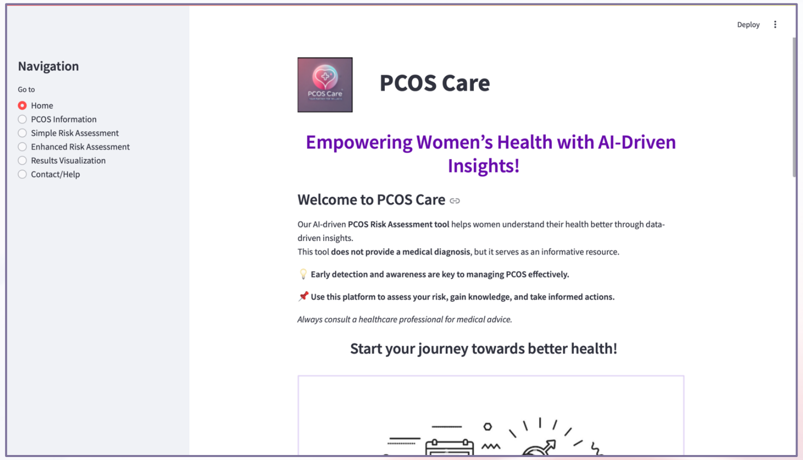
03. Initial Frontend Pages

- Basic UI developed using Python Streamlit.

- Pages include

1. Home page
2. PCOS Information*
3. Simple Risk Assessment*
4. Enhanced Risk Assessment*
5. Results Visualization*
6. Contact & Help*

** All pages can be viewed through the Github Repository*



PCOS Care

Empowering Women's Health with AI-Driven Insights!

Welcome to PCOS Care

Our AI-driven **PCOS Risk Assessment tool** helps women understand their health better through data-driven insights.
This tool **does not provide a medical diagnosis**, but it serves as an informative resource.

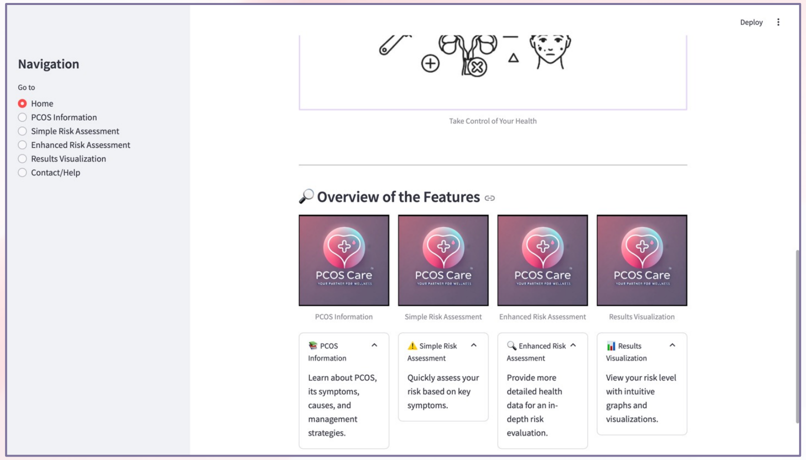
💡 **Early detection and awareness are key to managing PCOS effectively.**

📌 **Use this platform to assess your risk, gain knowledge, and take informed actions.**

Always consult a healthcare professional for medical advice.

Start your journey towards better health!

Frontend - Home page 1 (Self composed)



Take Control of Your Health

🔎 Overview of the Features

PCOS Information — Simple Risk Assessment — Enhanced Risk Assessment — Results Visualization

📚 PCOS Information: Learn about PCOS, its symptoms, causes, and management strategies.

⚠️ Simple Risk Assessment: Quickly assess your risk based on key symptoms.

🔍 Enhanced Risk Assessment: Provide more detailed health data for an in-depth risk evaluation.

📊 Results Visualization: View your risk level with intuitive graphs and visualizations.

Frontend - Home page 2 (Self composed)

Challenges Faced & Solutions

Challenge	Solution
Dataset having missing values	Handled by replacing with median or mode.
Dataset being imbalanced	Not relying solely on accuracy, checked other metrics such as precision, recall, etc. too.
Balancing recall vs. precision	Tuned model hyperparameters and focused on key features.
User entering invalid data	Implemented data validation.

Standout Achievements

Achieved High Accuracy

- Accuracy of 89.09% (simple risk) and 96.36% (enhanced risk), increasing its reliability.

Optimized Performance

- Successfully done hyperparameter tuning for both models.

Reduced False Negatives

- Reducing false negatives is crucial because a false negative means an actual PCOS-positive patient is wrongly predicted as negative, leading to missed diagnosis.

Model Evaluation after hyperparameter tuning

Assessment	Model	Accuracy	Precision	Recall	F1 Score	Specificity	FN
Simple	SVM	0.8909	0.9286	0.7222	0.8125	0.9730	5
Enhanced	Random Forest	0.9636	0.9444	0.9444	0.9444	0.9730	1

Prototype Overview

Key Functionalities

- User-friendly data input form for symptoms and history.
- ML-based PCOS risk prediction.
- Statistical analysis for result visualization.

Working Prototype Features

- Prediction Model
- UI Mock-ups
- Frontend pages

What It Demonstrates

A streamlined workflow for users to input health data, run PCOS predictions, and get prediction.

Hosting Details

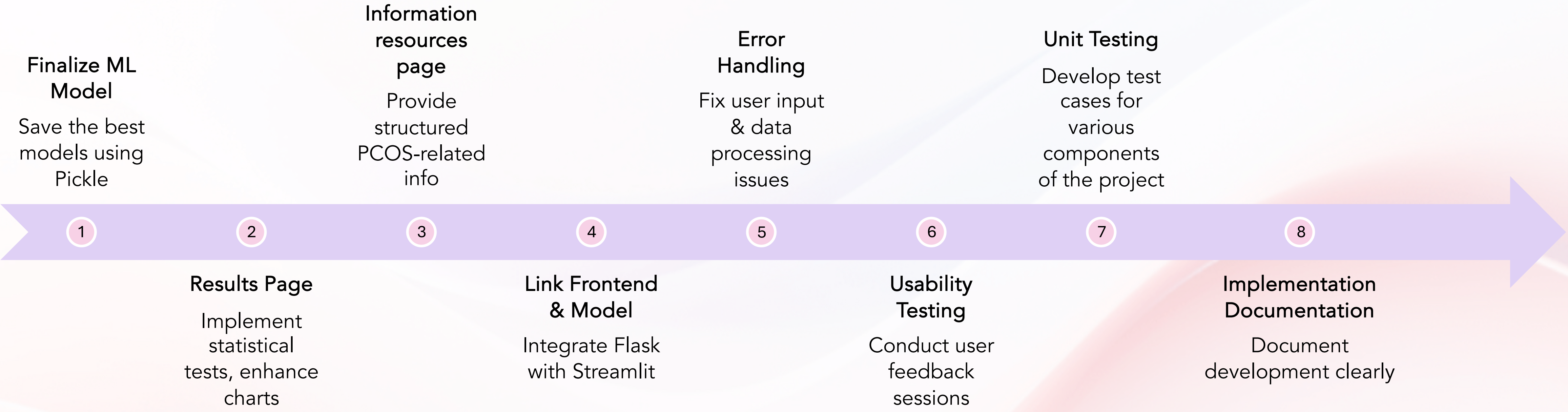
Currently running on localhost for development.

Video Demonstration Highlights

- Showcases user navigation, user data entry, ML model and other key frontend pages.
- Tools used

1. Jupyter Notebook (ML model)
2. Figma (UI mock-ups)
3. Python Streamlit (Frontend files)

Upcoming Milestones



1 **Finalize ML Model**

Save the best models using Pickle

2 **Results Page**

Implement statistical tests, enhance charts

3 **Information resources page**

Provide structured PCOS-related info

4 **Link Frontend & Model**

Integrate Flask with Streamlit

5 **Error Handling**

Fix user input & data processing issues

6 **Usability Testing**

Conduct user feedback sessions

7 **Unit Testing**

Develop test cases for various components of the project

8 **Implementation Documentation**

Document development clearly

Conclusion

DONE ✓

1. Successfully built the ML-based PCOS prediction system.
2. Improved model accuracy, reducing false negatives for better prediction.
3. Laid a solid foundation for development of frontend and future enhancements.

UPCOMING

1. Finalizing pending features, including result visualization.
2. Addressing hosting challenges while focusing on security and scalability.
3. Ensuring a fully functional and optimized web application by the deadline.

References

Adla, Y. A. et al. (2021). Automated Detection of Polycystic Ovary Syndrome Using Machine Learning Techniques. 2021 Sixth International Conference on Advances in Biomedical Engineering (ICABME). Available from http://dx.doi.org/10.1109/ICABME53305.2021.9604905 [Accessed 6 October 2024].

Battiste, J. (2024). Women with painful condition 'ignored' by doctors. BBC News, 3 June. Available from https://www.bbc.co.uk/news/articles/clwwgl98dg0o [Accessed 8 January 2025].

OpenAI. (2025). App Logo Image for PCOS Care, generated using DALL·E via ChatGPT. [Accessed 12 January 2025].

Thank you