Web Application for Diagnosing PCOS Patients Using Machine Learning

Student: Rachel Naadia Cooray (19564441)

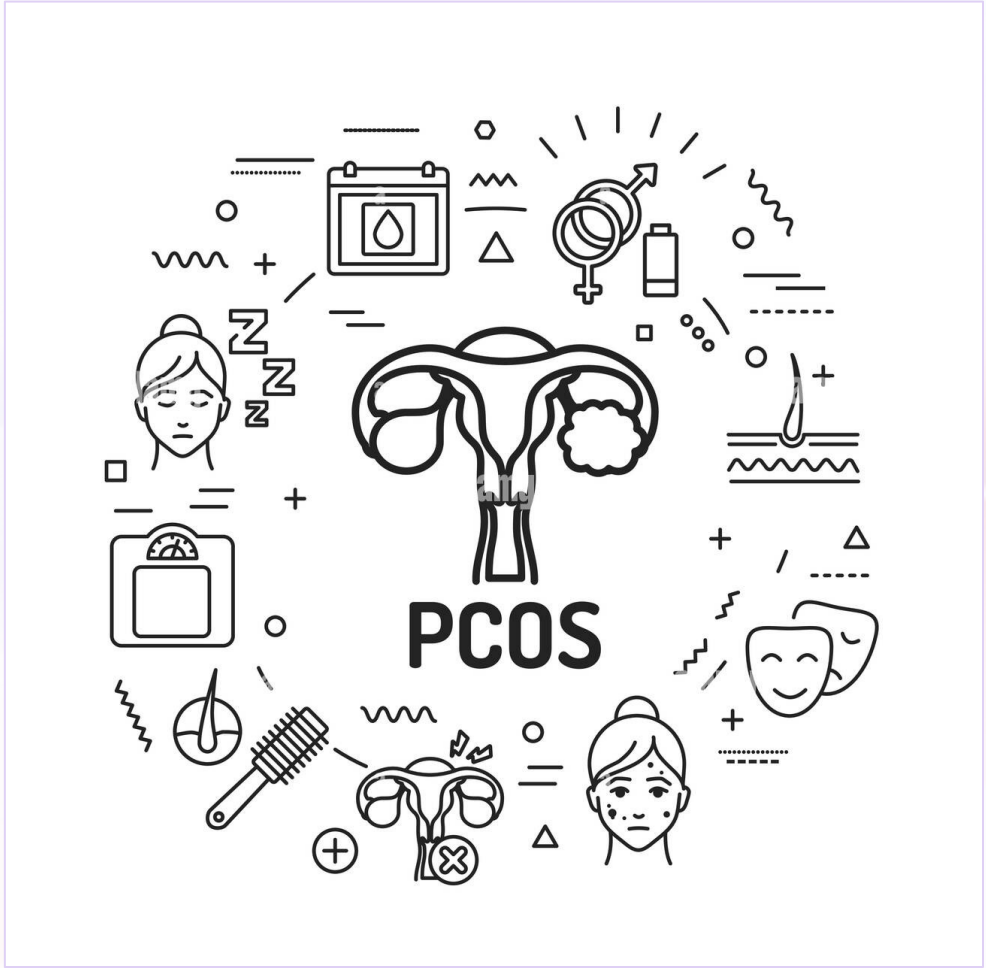
Supervisor: Dr Malarvizhi Kaniappan Chinnathai

Degree: BSc (Hons) Computer Science

14th November 2024

Content

- Problem Description
- Aims and Objectives
- Background Review
- Comparison with Other Systems
- Tools and Technologies
- Key Features
- Use Cases
- Diagrams
- Requirements Elicitation Techniques
- Requirements Elicitation
- Project Timeline & Milestones
- References



Problem Description

OVERVIEW

Polycystic Ovary Syndrome (PCOS) is a common endocrine disorder affecting women of reproductive age of which 70% are undiagnosed (Adla et al., 2021).

Undiagnosed PCOS can lead to severe health issues, including infertility, metabolic disorders, and mental health challenges (Gandhi, Patel and Dave, 2024).

LIMITATIONS

Diagnosis often requires in-person visits, which are time consuming and costly.

Many cases go undiagnosed due to a lack of accessible diagnosis tools. (Adla et al., 2021)

SOLUTION

A reliable, user-friendly tool for assessing PCOS risk is needed, enabling women to evaluate their risk conveniently.

Early risk assessment empowers users to seek timely intervention, improving health outcomes and quality of life.

Aim and Objectives

AIM

To develop a web application that predicts the likelihood of PCOS and provides educational resources to users.

OBJECTIVES

1. Enable users to input health data and symptoms.
2. Provide accurate PCOS predictions in a user friendly manner.
3. Display educational resources on PCOS.

Background Review

Algorithms & Accuracy

- Machine learning techniques like Random Forest, Logistic Regression, Decision Trees Priyadharshini et al. (2024), and SVM have been applied in PCOS prediction
- Random Forest achieved an accuracy of 83.48% (Gandhi, Patel and Dave, 2024) and 93.25% (Tiwari et al., 2022) in separate studies.

Feature Selection

- Significant features were identified using methods like the filter approach and ANOVA, helping improve prediction accuracy by focusing on essential attributes (Adla et al., 2021).

Model Limitations

- Challenges include balancing recall and precision, with some models showing high accuracy but weaker recall, missing some true positive cases (Adla et al., 2021).

Dataset Needs

- Studies emphasize the necessity of larger, high-quality, standardized datasets to improve model accuracy, and datasets from diverse populations (Radhakrishnan et al., 2023).

Future Directions

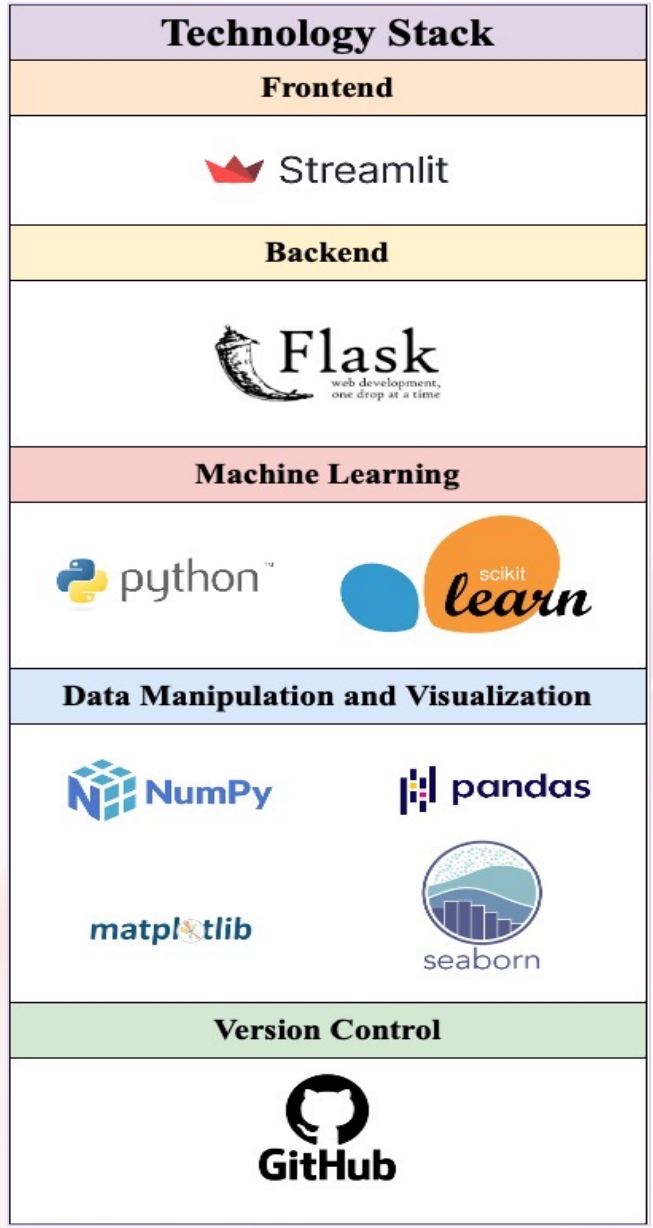
- Improved feature selection, advanced algorithms, and more inclusive datasets are needed to generalize models effectively and increase adoption of AI in PCOS diagnosis.

Existing systems for PCOS

Product	PCOS Information	User data input	PCOS Prediction	Results visualisation
NHS Website	Yes	No	No	No
Flo App	Yes	Yes	No	No
PCOS App	Yes	No	No	No
PCOS Tracker	Yes	No	No	No
Project to be implemented	**Yes**	**Yes**	**Yes**	**Yes**

Tools & Technologies

Tool	Resource Type
Laptop/PC with 8GB RAM or more	Hardware
Streamlit	Frontend Framework
Flask	Backend Framework
Python	Programming Language
scikit-learn	Machine Learning Library
Pandas and NumPy	Data Manipulation
Matplotlib and Seaborn	Visualization Library
GitHub	Version Control



Tech stack (Self composed)

Key Features

Data Input Form

User-friendly, guided form for accurate data entry.

01

Information Page

Detailed section on PCOS symptoms, causes, and management.

02

Prediction Model

Choose the best classification model based on evaluation.

03

Result Visualization

Graphs and charts showing prediction results.

04

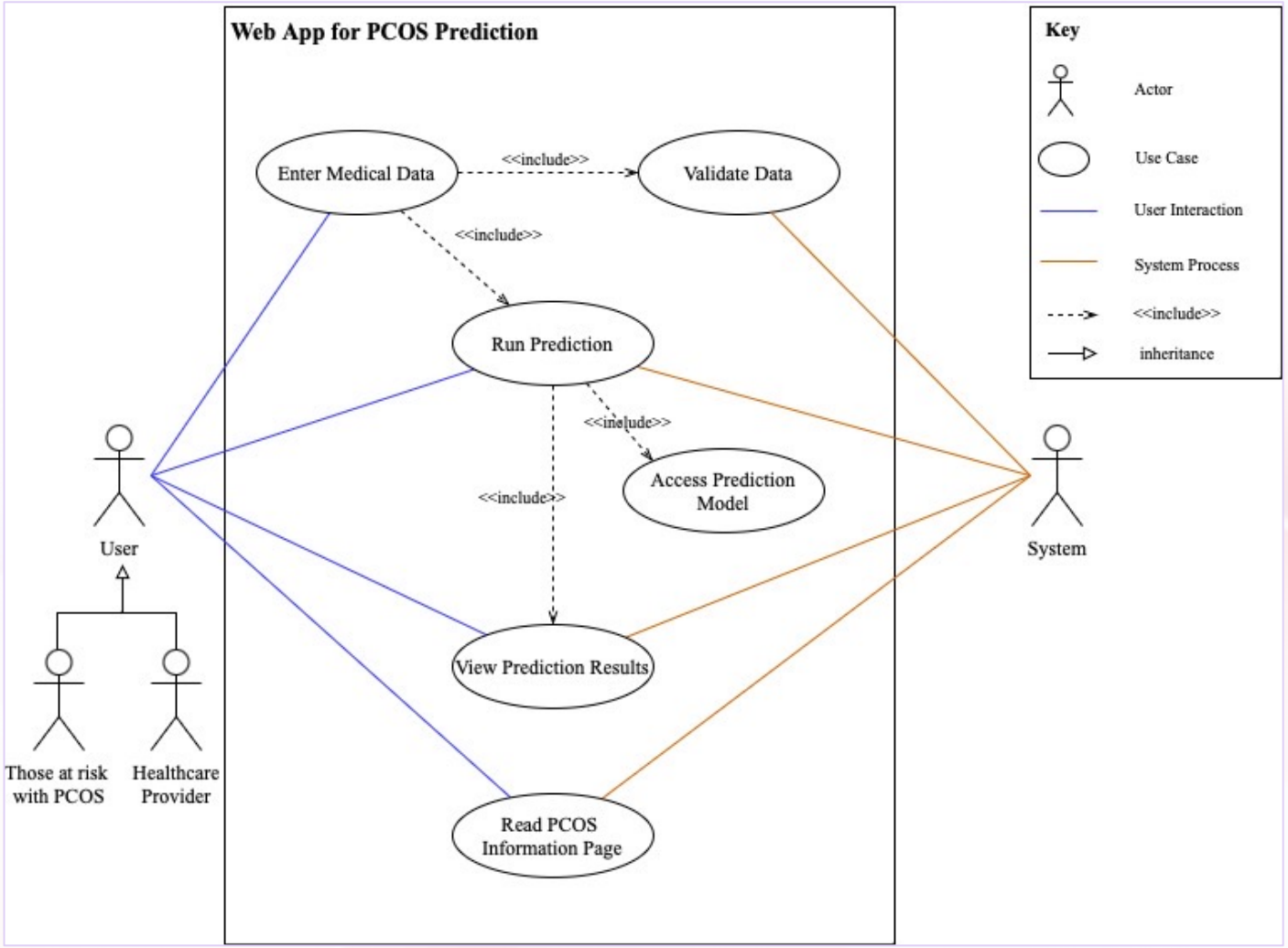
Use cases

Actors

1. Those at risk with PCOS
2. Healthcare providers
3. System

Use cases

1. Enter Medical Data
2. Run Prediction
3. View Prediction Results
4. Read Information on PCOS

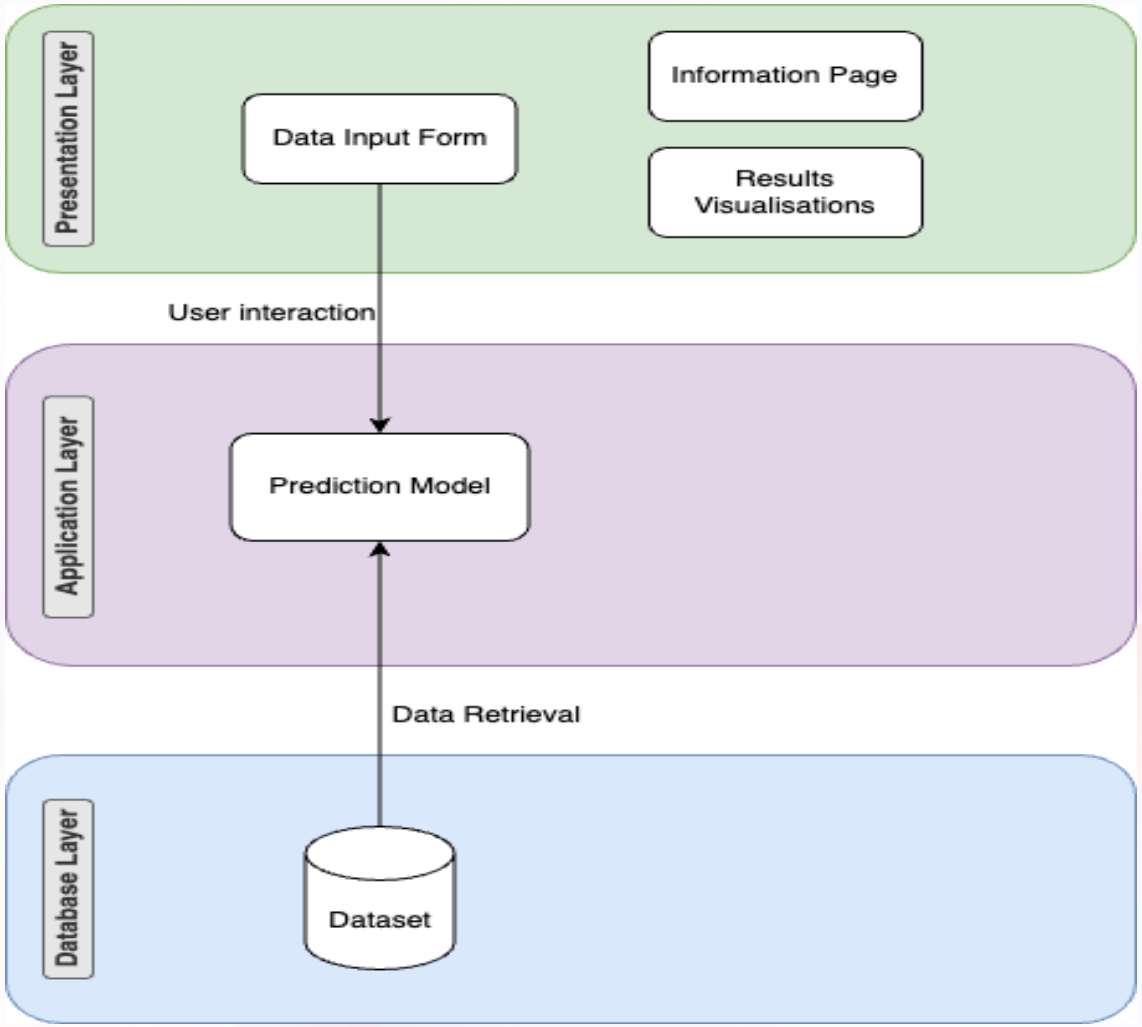


Use case Diagram (Self composed)

Diagrams

1. System Architecture Diagram

Helps understand the interactions among entities and system.

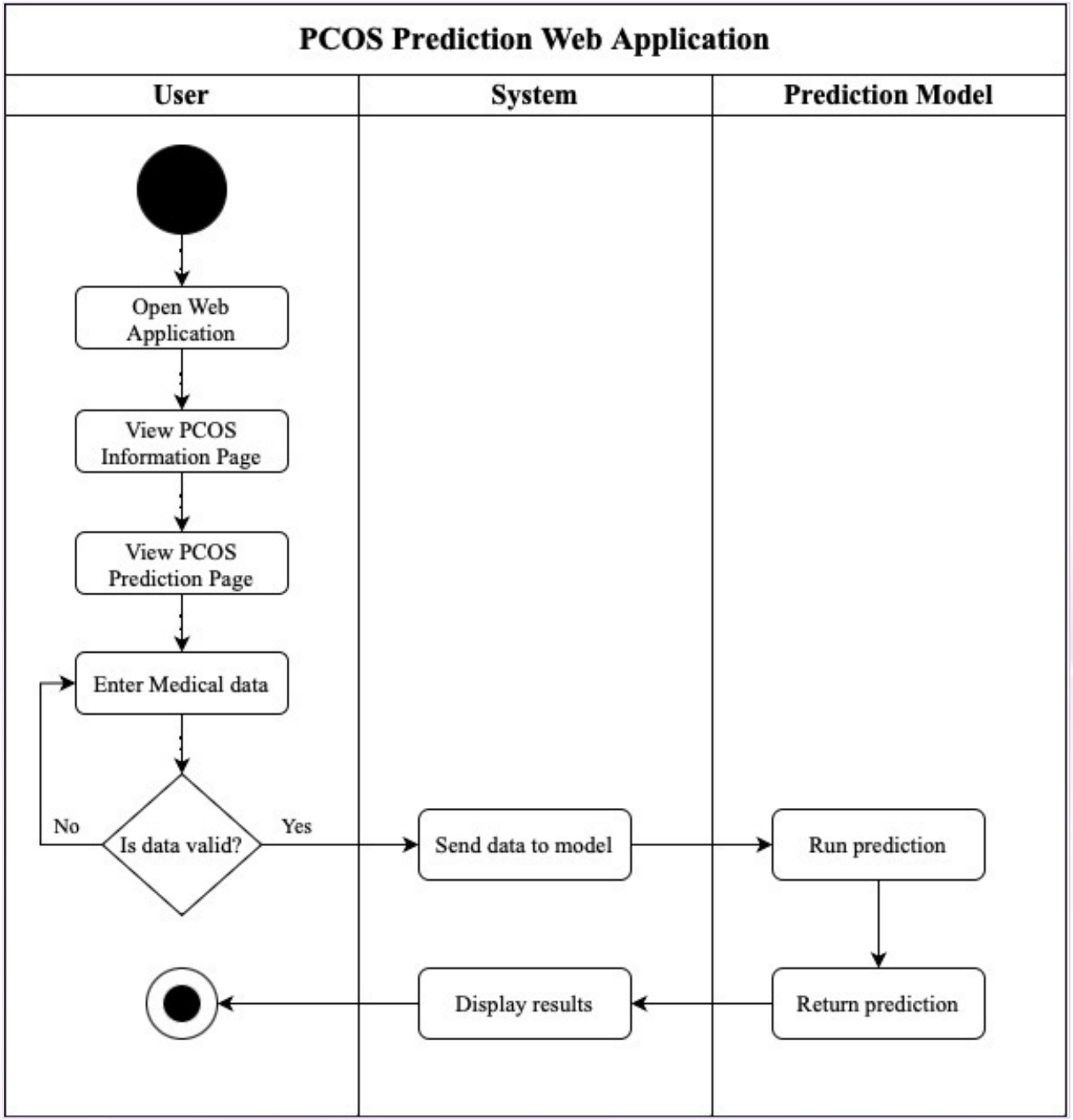


System Architecture Diagram (Self composed)

Diagrams

2. Activity Diagram

The data flow through different steps in a system.

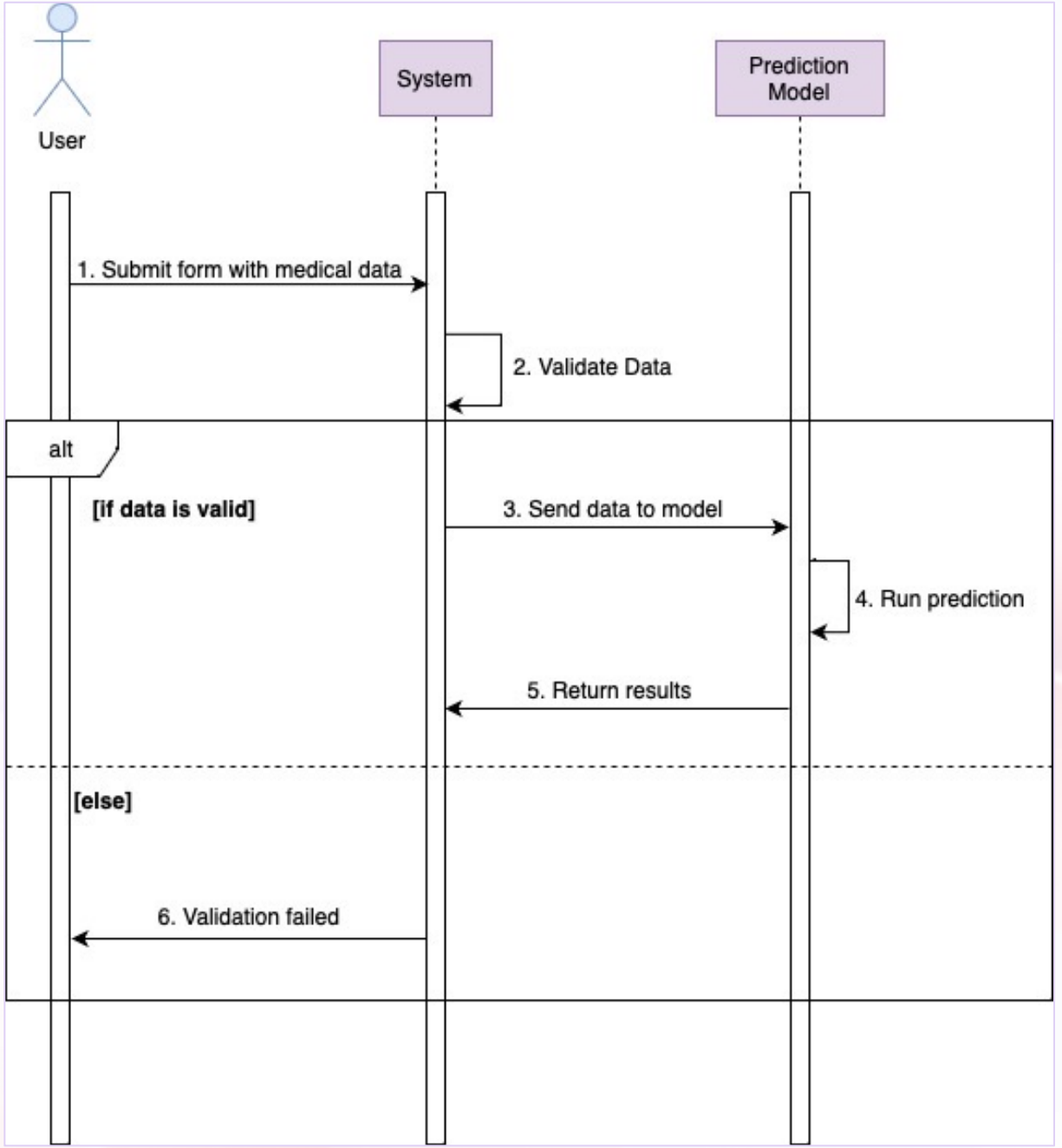


Activity Diagram (Self composed)

Diagrams

3. Sequence Diagram

How the system objects interact with one another to complete a task.



Sequence Diagram (Self composed)

Requirements Elicitation Techniques

Literature Review

• Reviewing and understanding existing PCOS studies, their features and limitations.

Dataset Analysis

• Identifying relevant data from the Kaggle PCOS dataset (Kottarathil, 2021).

Questionnaires

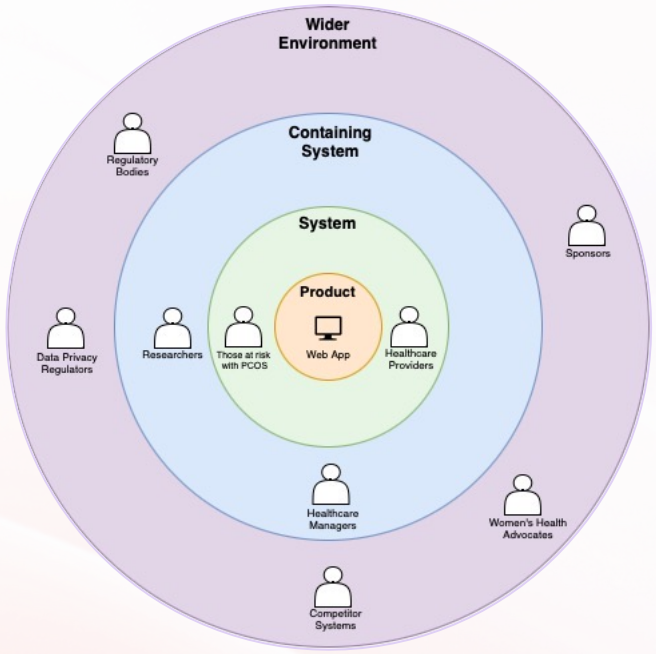
• Gathering ideas from users on their expectations of the product.

Brainstorming

• Creative feature ideation.

Requirements Elicitation

Stakeholders analysis



Onion Model (Self composed)

Indentification of system requiremnets

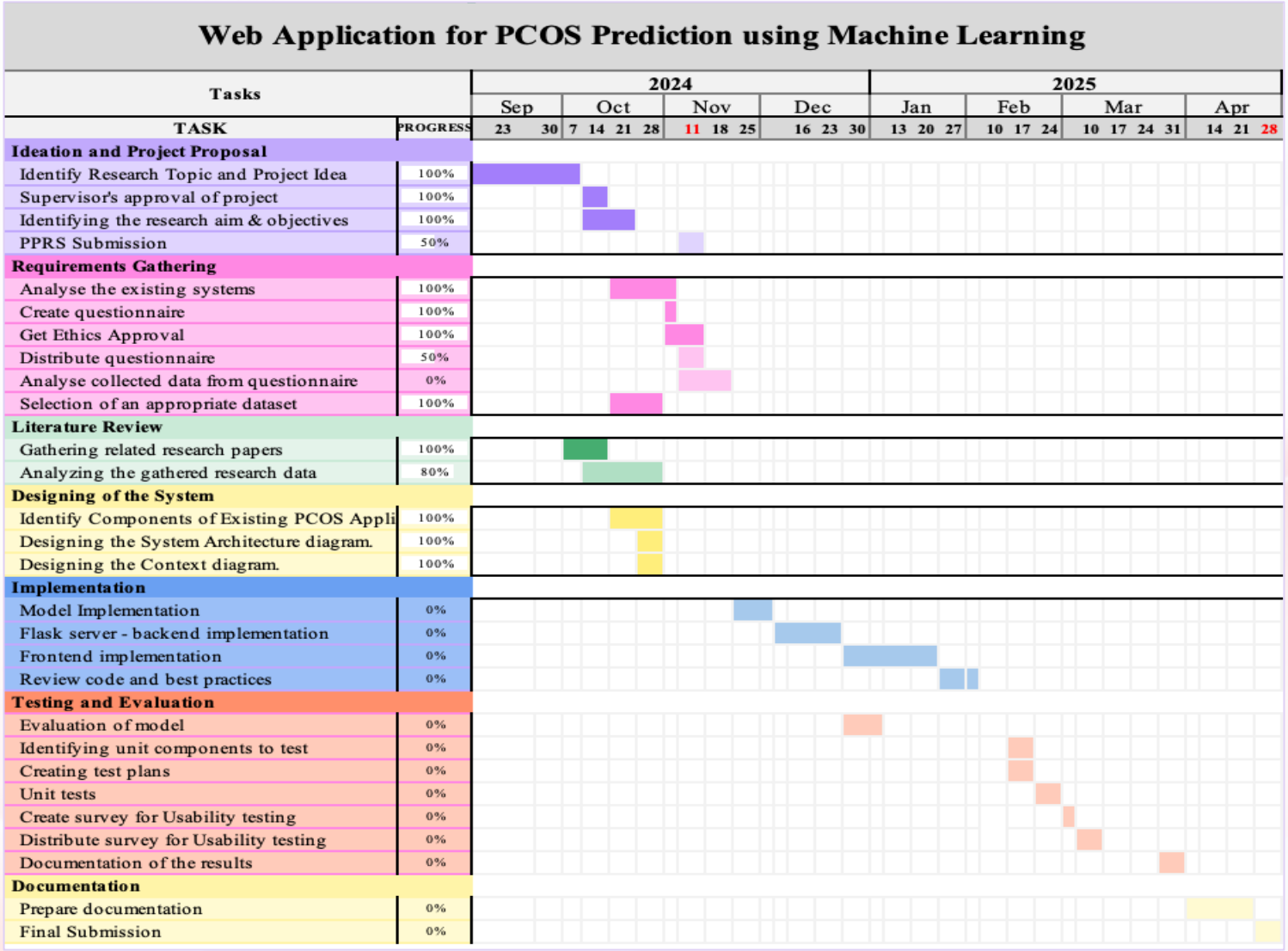
Functional Requirements:

- Data Input
- Prediction Model
- Results Display
- Information Resources

Non-functional Requirements:

- Performance
- Usability
- Security
- Compatibility

Project Timeline & Milestones



Gantt Chart (Self composed)

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].

Gandhi, D., Patel, B. and Dave, N. (2024). PCOS Detection using Machine Learning Algorithms. *Indian Journal of Natural Science*, 14 (81). Available from https://www.researchgate.net/publication/381456665_PCOS_Detection_using_Machine_Learning_Algorithms [Accessed 8 October 2024].

Hdaib, D. et al. (2022). Detection of Polycystic Ovary Syndrome (PCOS) Using Machine Learning Algorithms. *5th International Conference on Engineering Technology and its Applications 2022.* Available from https://doi.org/10.1109/IICETA54559.2022.9888677 [Accessed 6 October 2024].

Kottarathil, P. (2021). Polycystic ovary syndrome (PCOS). Available from https://www.kaggle.com/datasets/prasoonkottarathil/polycystic-ovary-syndrome-pcos/data [Accessed 6 October 2024].

NHS. (2022). Polycystic ovary syndrome. Available from https://www.nhs.uk/conditions/polycystic-ovary-syndrome-pcos/ [Accessed 6 October 2024].

Priyadharshini, M. et al. (2024). PCOS Disease Prediction Using Machine Learning Algorithms. *International Research Journal on Advanced Engineering Hub (IRJAEH)*, 2, 651—655. Available from http://dx.doi.org/10.47392/IRJAEH.2024.0094 [Accessed 6 October 2024].

Radhakrishnan, R. et al. (2023). Computerized Diagnosis of Polycystic Ovary Syndrome Using Machine Learning and Swarm Intelligence Techniques. Available from https://doi.org/10.21203/rs.3.rs-2027767/v2 [Accessed 6 October 2024].

Tiwari, S. et al. (2022). SPOSDS: A smart Polycystic Ovary Syndrome diagnostic system using machine learning. *Expert Systems with Applications 203*, 1. Available from http://dx.doi.org/10.1016/j.eswa.2022.117592 [Accessed 6 October 2024].

Thank you