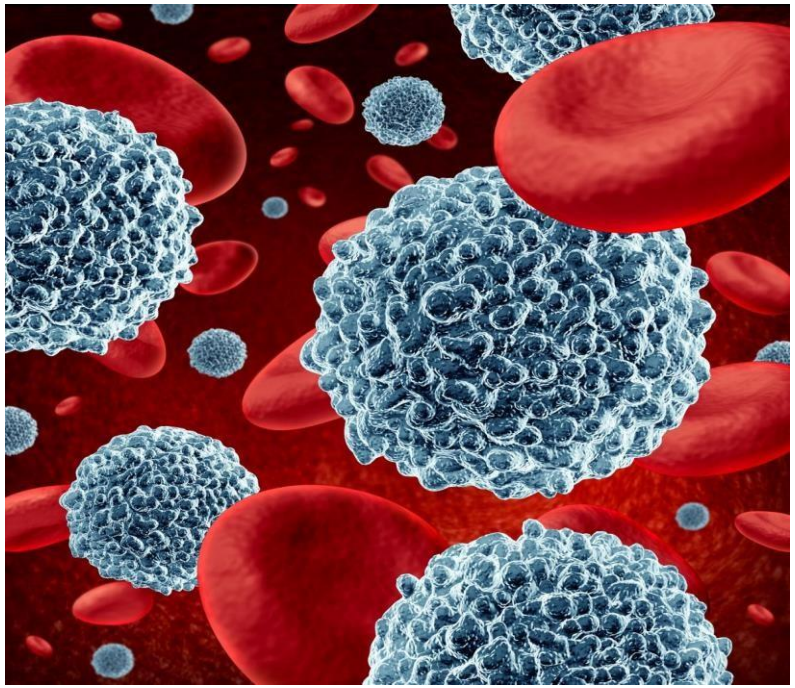
Anemia level prediction in children

2018 Nigeria Health Demographic Survey

Team Members

- Ruth Kitasi
- Agatha Nyambati
- Joselyne Apiyo
- Cecilia Ngunjiri
- John Mbego
- Leonard Koyio



Overview

Anemia is a major health issue among children aged 0-59 months in Nigeria, impairing growth, cognition, and immunity. It is driven by poor nutrition, frequent infections, and limited healthcare access, with socioeconomic factors like low income, limited education, and poor sanitation worsening its prevalence.

Addressing anemia is vital for improving child health and supporting Nigeria's development.

Problem Statement

Many children in Nigeria suffer from undetected and untreated anemia due to limited healthcare access, high diagnostic costs, and low public awareness.

This project aims to develop a model that predicts anemia severity based on socioeconomic factors, enabling the identification of high-risk children and supporting targeted health interventions.

Objectives

1. Develop a classification model for Anemia severity in children

2. Evaluate the model's predictive accuracy for effective Public Health Use

3. Identify key socioeconomic determinants influencing Anemia severity

4. Provide Actionable Insights for Targeted Interventions.

5. Long-Term Policy Development for Anemia Reduction

Data Source

The data source for this project was the Nigeria Demographic and Health survey which is a comprehensive survey designed to provide essential data on the health and demographic status of Nigeria's population.

The data is obtained from the 2018 Nigeria Demographic and Health Surveys to answer research questions about the effect of mothers' age and other socioeconomic factors on children aged 0-59 months anemia level.

Data Limitations

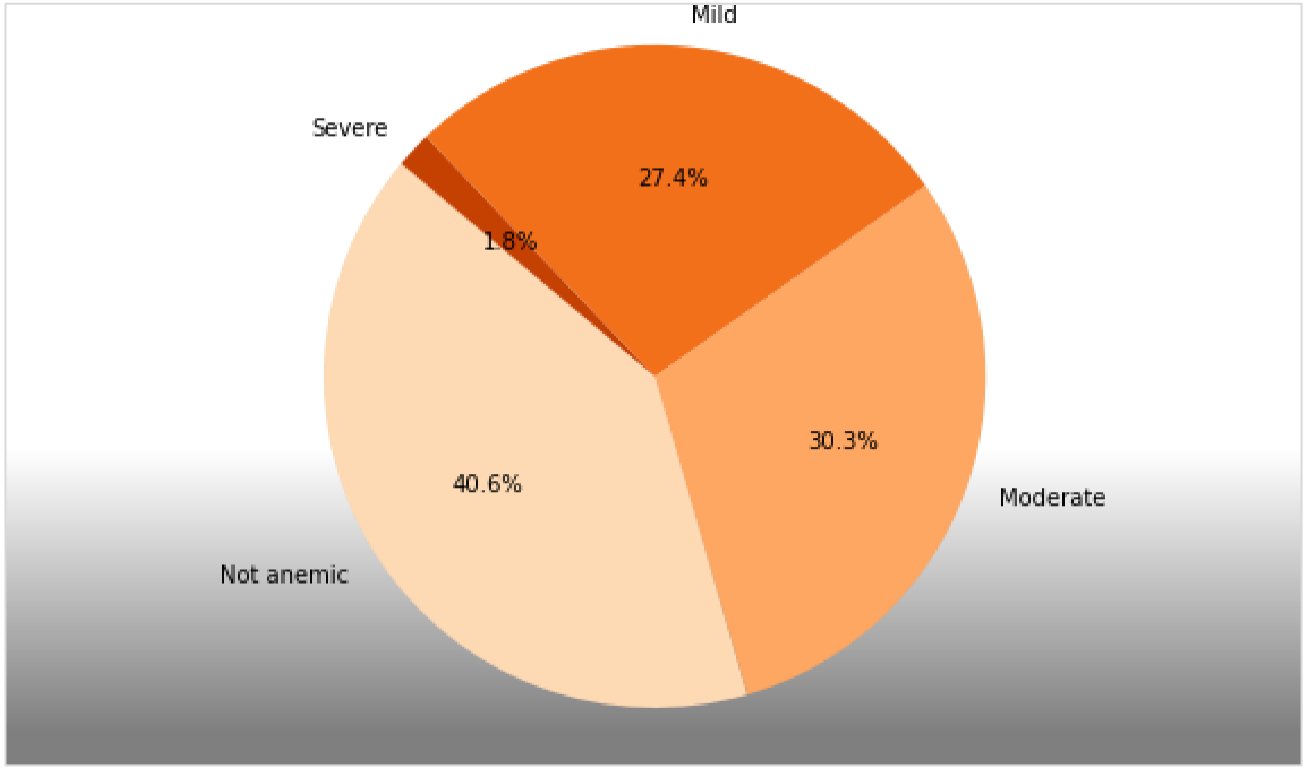
- Data quality and completeness where missing values, inconsistencies, or errors in the dataset result in misinterpretations.
- Limited sample sizes or non-representative samples can affect the generalizability of the findings.
- Socioeconomic factors may be underreported or inaccurately recorded, especially in informal settlements or among marginalized populations
- Temporal factors which may change over time due to policy changes, economic fluctuations, or health interventions leading to outdated conclusions.
- Cultural biases or reluctance to disclose certain behaviors can affect the reliability of this information.

Important Features

- ✓ Anemia Level
- ✓ Hemoglobin level
- ✓ Educational level
- ✓ Area(urban/rural)
- ✓ Wealth Index
- ✓ Age group

Anemia level distribution

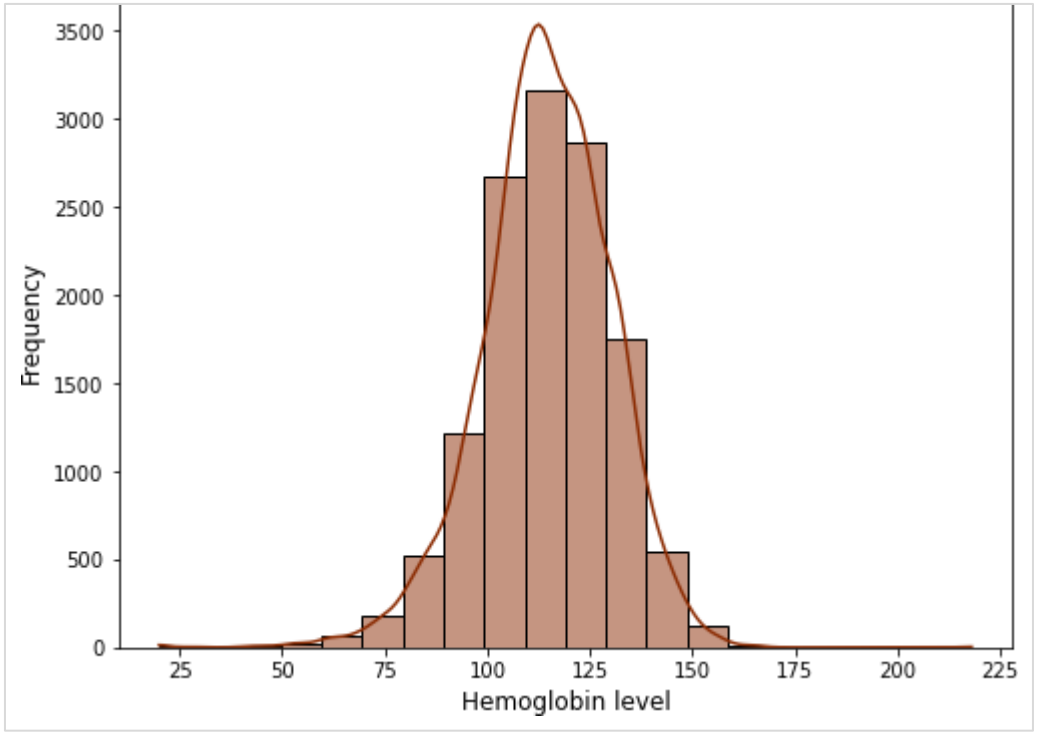
- This pie chart shows that most children (60%) have moderate, mild, or severe anemia.
- The variation in levels indicates that certain types of anemia are more prevalent than others



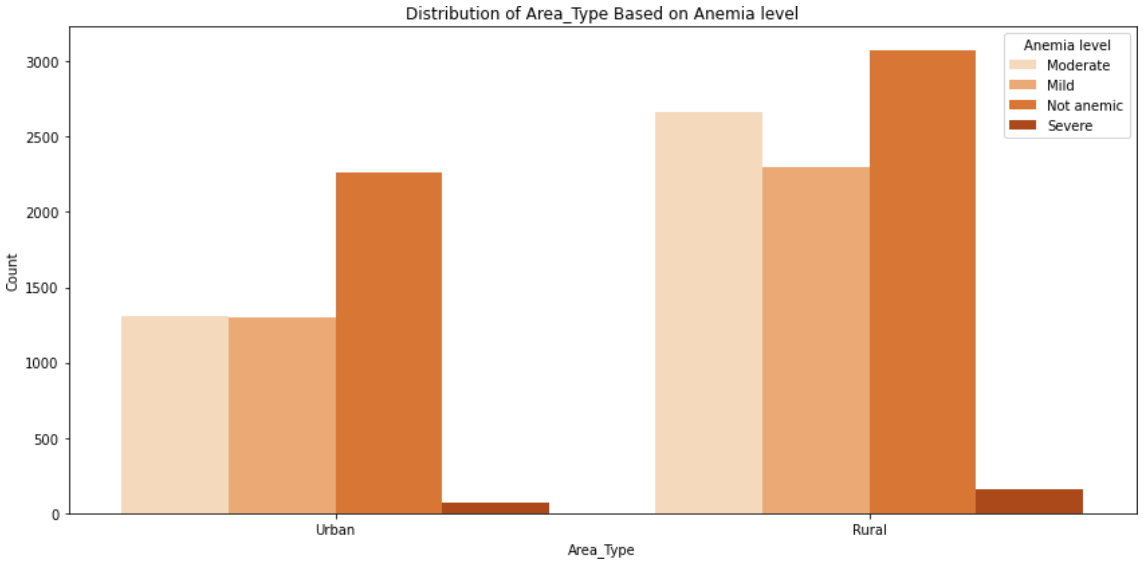
Hemoglobin level distribution

- The graph shows a normal distribution of hemoglobin levels among the respondents, with most values concentrated between 95 and 125 levels.

- This shape implies that hemoglobin levels are normally distributed among the respondents after adjusting for altitude.



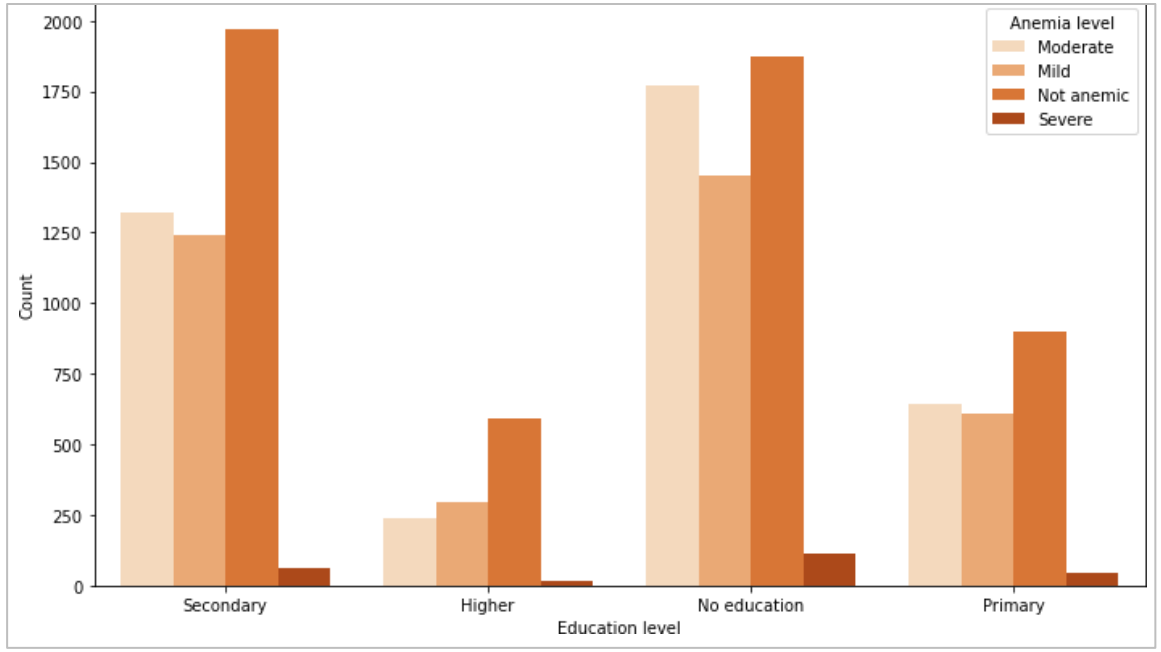
Explanatory Data Analysis



Distribution of Area type based on anemia level in children

- The distribution of anemia levels differs between rural and urban areas, with rural areas showing a higher prevalence of moderate anemia.
- This suggests potential disparities in socioeconomic factors or healthcare access between these settings.

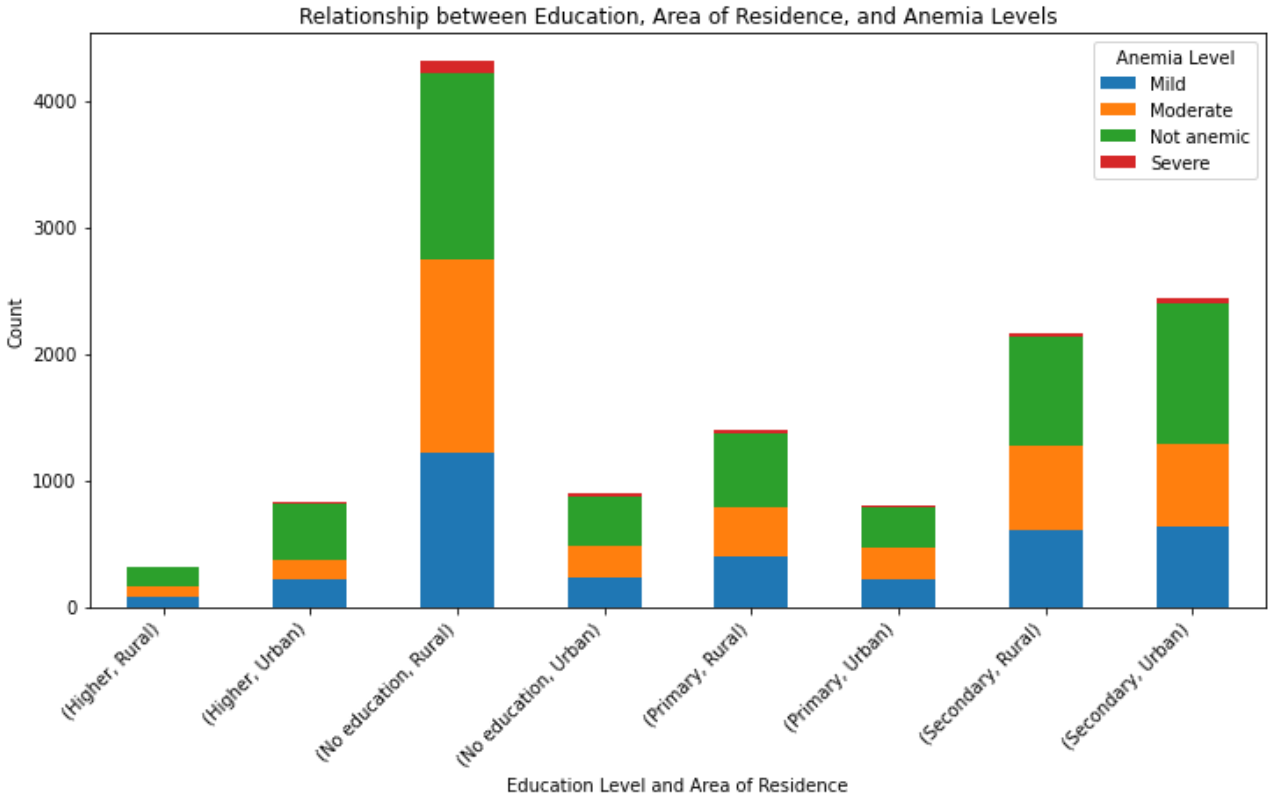
Explanatory Data Analysis



Distribution of maternal education level based on anemia level in children

•The analysis shows a clear relationship between education level and anemia severity. Children whose parents have no education have the highest prevalence of moderate anemia, suggesting that lack of education is associated with poorer anemia outcomes.

Explanatory Data Analysis



Distribution of area and level of education based on anemia level in children

The combination of education and area of residence highly influence anemia risk. For example, children of individuals with low education in rural areas have a higher risk of anemia compared to those in urban areas with higher education.

Modelling & Evaluation

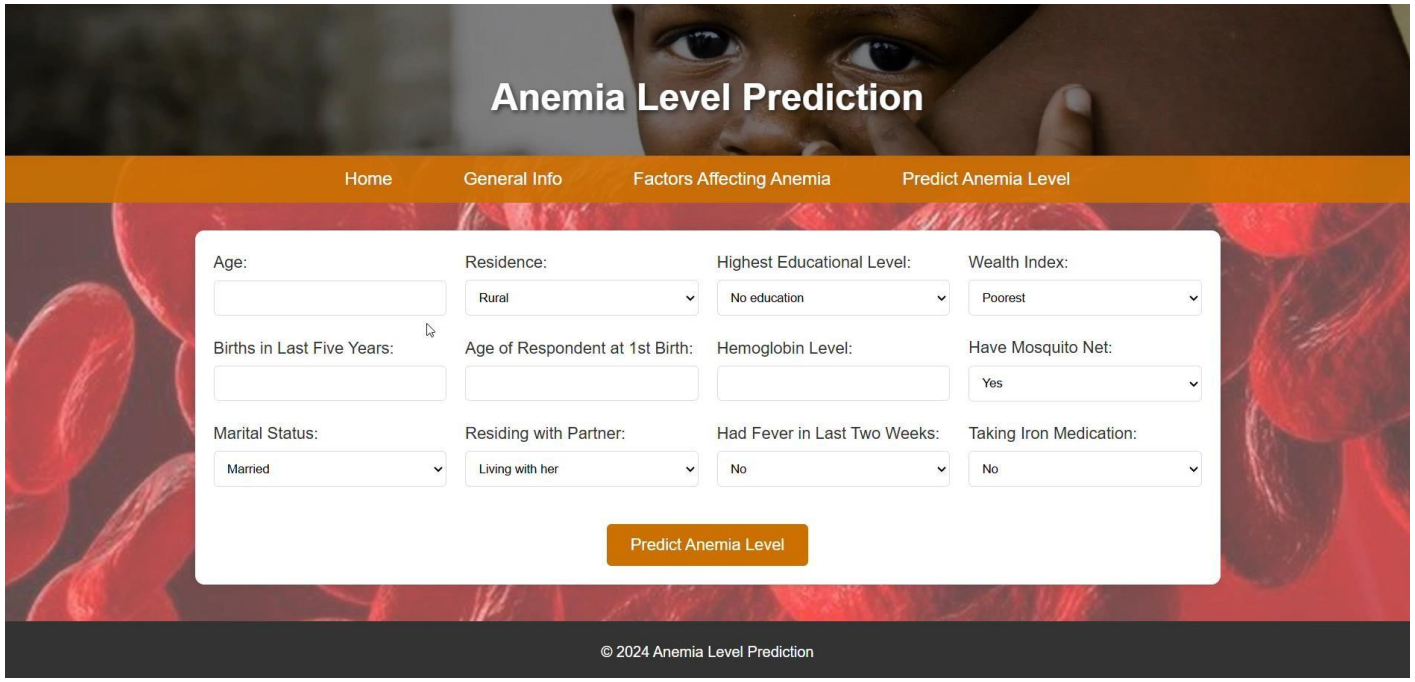
Metrics	Logical reg	Decision tree	Random forest	KNN
Accuracy	0.88	0.93	0.95	0.65
Precision	0.89	0.93	0.95	0.69
Recall	0.88	0.93	0.95	0.67
F1 Score	0.88	0.93	0.95	0.67

The baseline model, Logistic Regression, achieved an accuracy score of 88%. Among the four models tested, the Random Forest model performed the best, initially reaching an accuracy score of 95%.

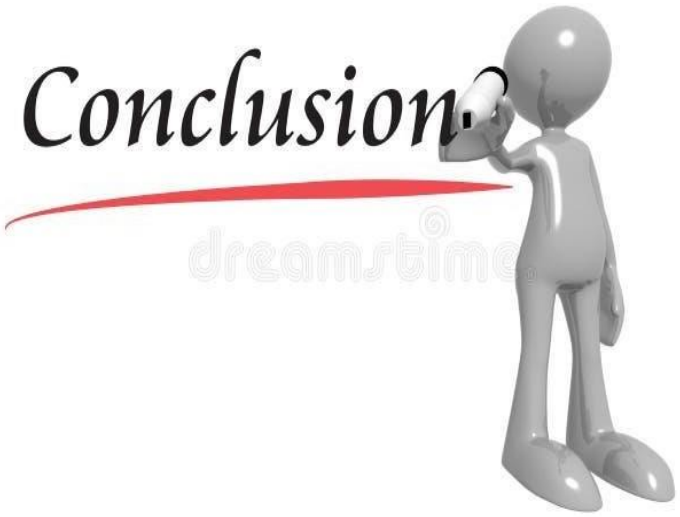
Best Model tunned, improving its accuracy to 96%, making it a strong candidate for deployment

Model Deployment

We developed our anemia-level prediction application using Flask, a lightweight and powerful web framework, and successfully deployed it on Render, ensuring a seamless and accessible user experience



Live link: Anemia level prediction application



- ✓ The project successfully developed a classification model using Python libraries to predict the severity of anemia among children under 59 months.
- ✓ Model Accuracy and Reliability: The model has undergone thorough testing and evaluation, demonstrating its ability to accurately identify the socio-economic factors influencing anemia.
- ✓ The impact of socio-economic factors such as wealth index is associated with anemia severity, and these findings emphasize the need for targeted interventions in communities with lower socio-economic status.
- ✓ The project's findings and the model’s application in real-world settings contribute to policy development/recommendations by offering reliable data highlighting socio-economic factors associated with anemia.
- ✓ Practical Application: A user-friendly web application has been developed, allowing users to input data and receive personalized predictions of anemia and tailored recommendations. This makes findings accessible for public health decision-making and individual-level intervention.

Recommendations

1. Integrate Nutritional & Socio-Economic Interventions: Combine nutritional supplements (e.g., iron, folic acid) with socio-economic initiatives to address anemia's root causes.

2. Develop Geo-Mapping Tools: Visualize anemia prevalence to identify high-risk regions and direct resources effectively for targeted interventions.

3. Integrate Predictive Models into Hospital Dashboards: Enhance risk assessment and patient monitoring in resource-limited settings for early detection and timely treatment.

4. Prioritize Hemoglobin Testing: Implement widespread screening programs, focusing on young children to improve diagnosis and intervention.

Next steps

•Deploy automated surveys to gather user feedback and measure improvements in anemia management.

•Integrate current and diverse datasets to ensure the model reflects the latest trends

•Enhance the model to include predictions for other related conditions, such as malnutrition and malaria, which often co-occur with anemia.

- *Special thank you to our technical mentors and the group 15 members who have made this project a success*