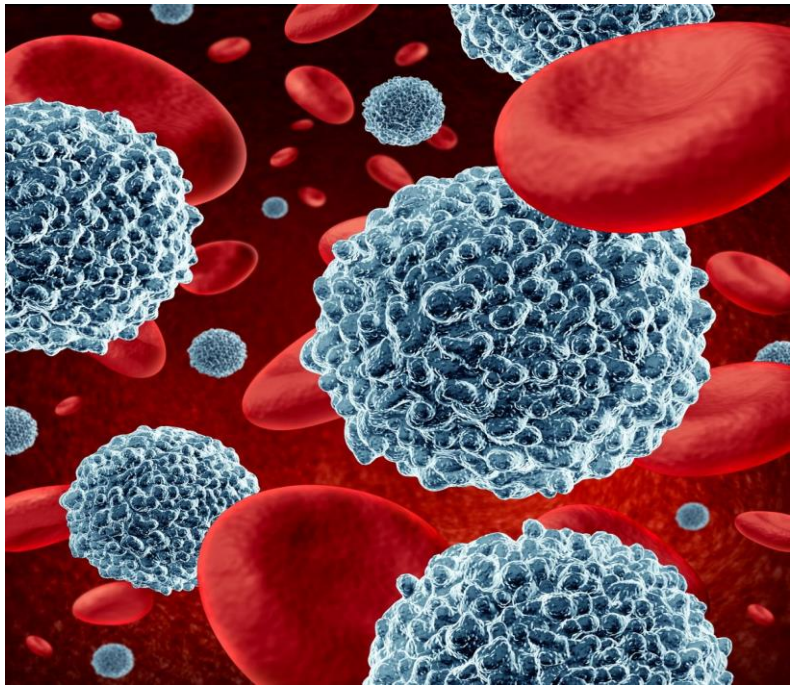
ANEMIA LEVEL PREDICTION IN CHILDREN

Team Members

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



Business Understanding

Anemia is a significant health issue, especially among Nigerian children under the age of five This condition hampers their growth, cognitive development, and immune health, driven primarily by poor nutrition, frequent infections, and limited healthcare access.

Socioeconomic factors, such as low household income, limited education, and lack of proper sanitation, further contribute to its high prevalence. Tackling anemia is essential, not only to improve children’s health outcomes but also for Nigeria's broader national development, given the long-term impacts on physical and cognitive well-being.

Problem Statement

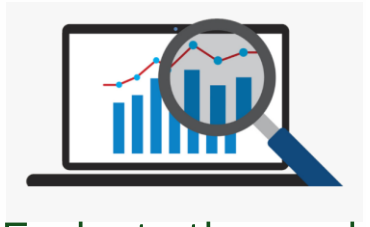
Many children in Nigeria suffer from undetected and untreated anemia due to limited healthcare access, high diagnostic costs, and low public awareness.

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. Identify key socioeconomic determinants influencing Anemia severity



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

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.

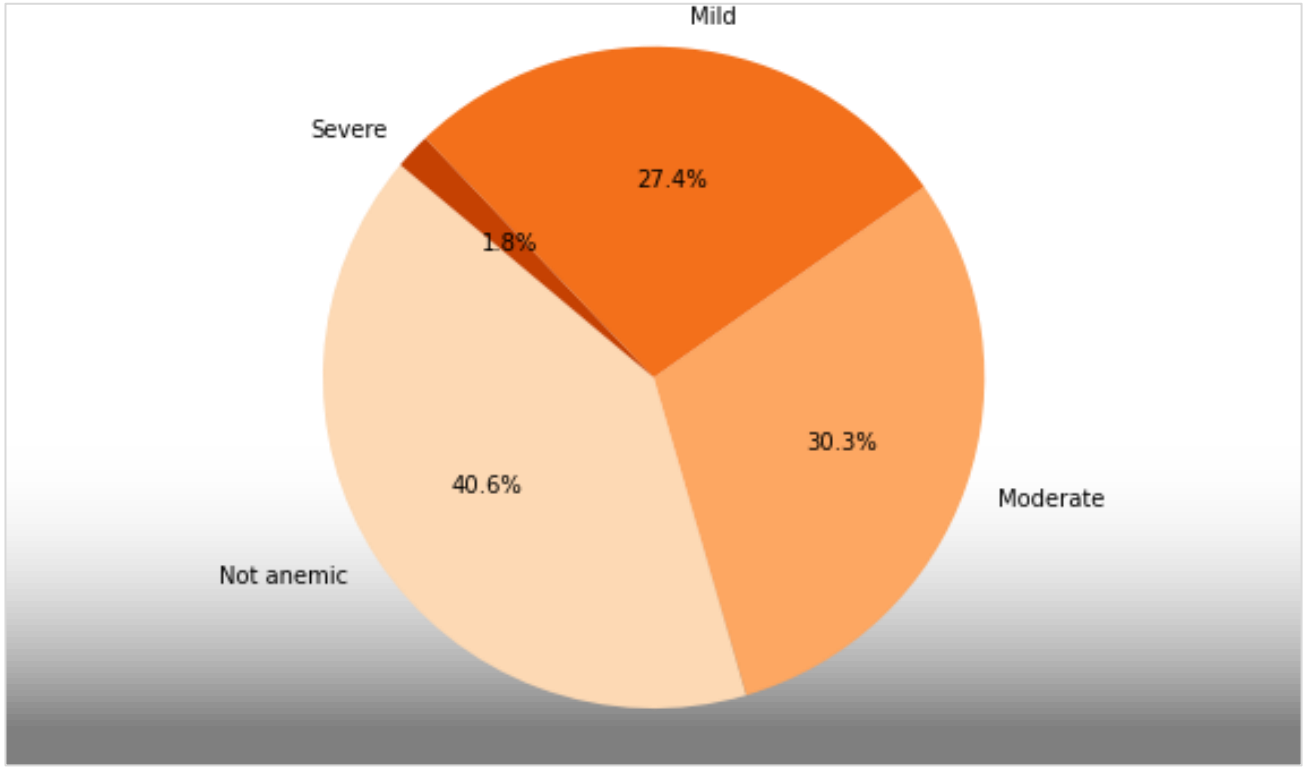
Important Features

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

Graphical representation of Important features

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

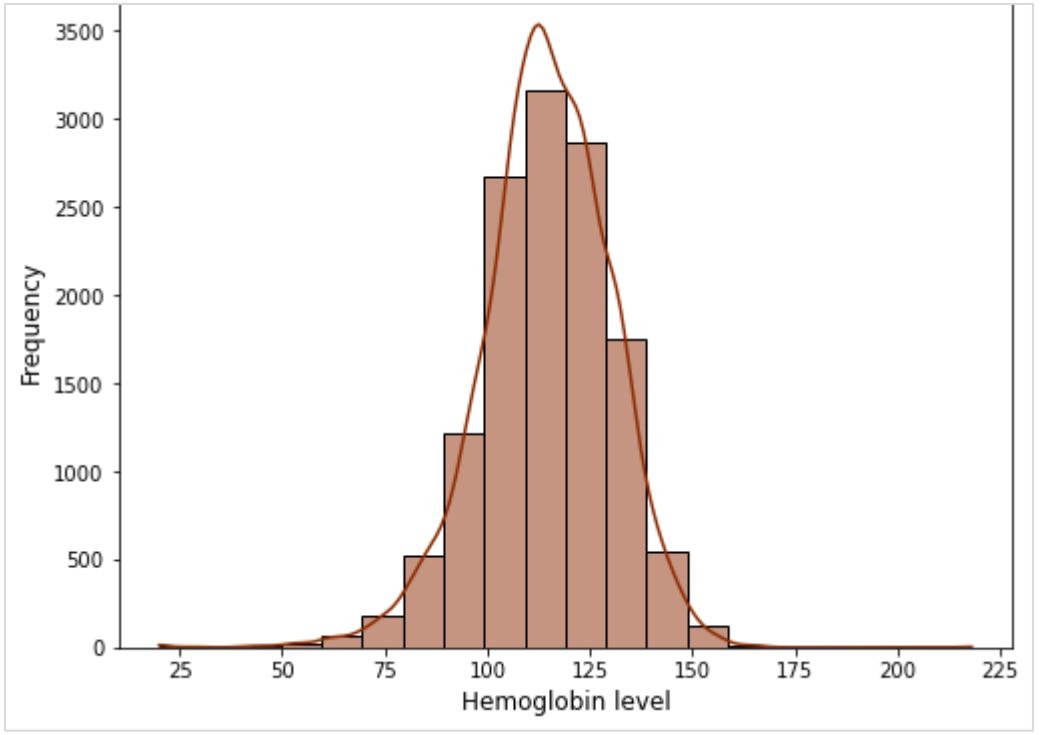


Graphical representation of Important features

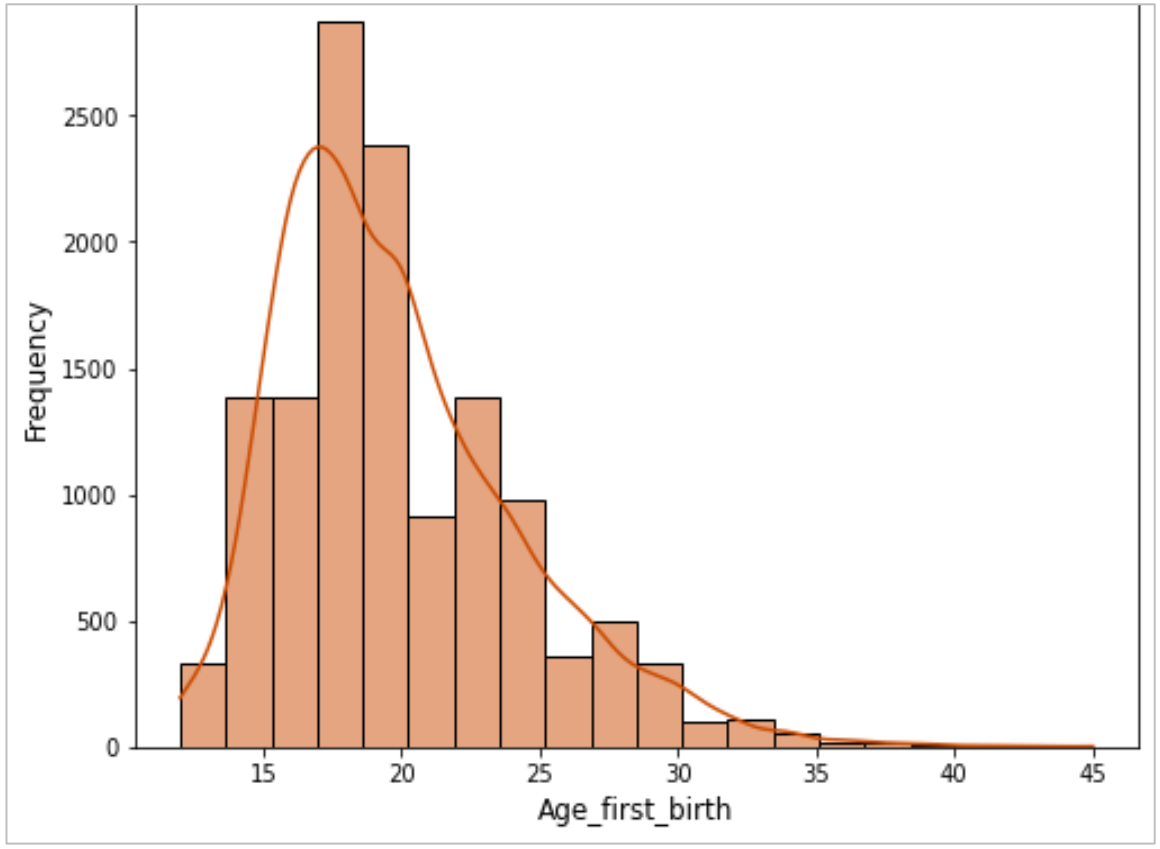
Hemoglobin level distribution

- 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.



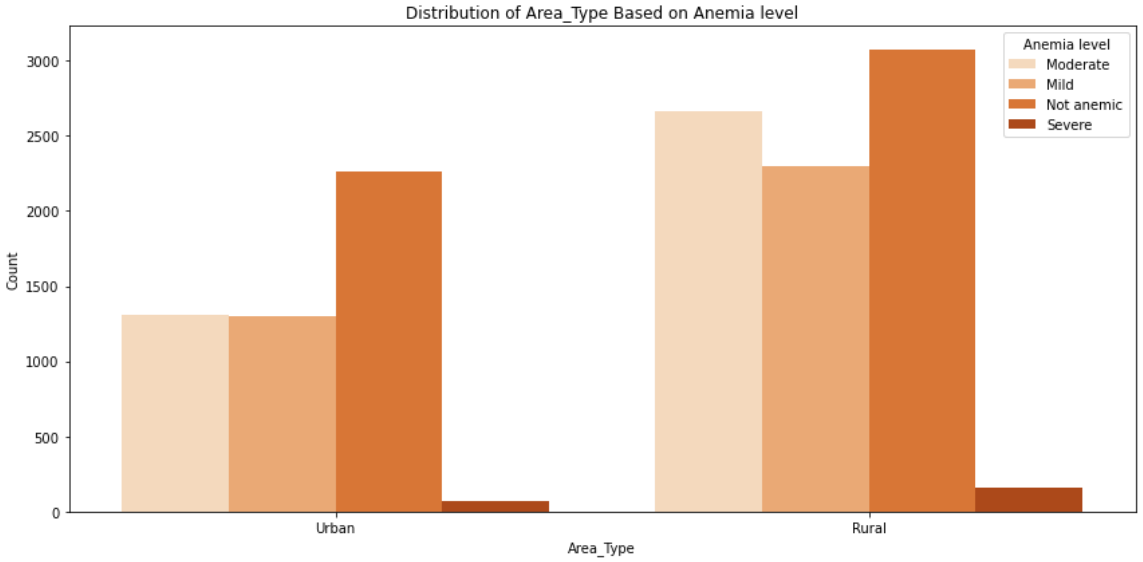
Graphical representation of Important features



Distribution of maternal age group at first birth

- The graph shows that most women in this dataset had their first child at a young age, with a peak between 15 and 25 years. This indicates a high prevalence of early childbearing.

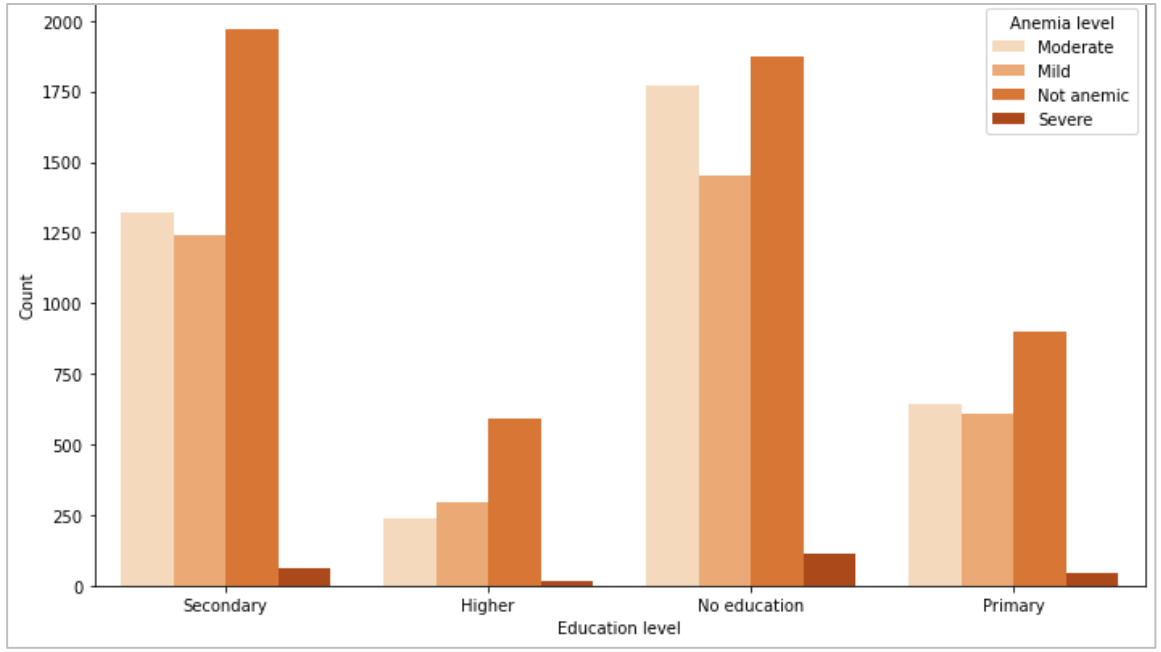
Graphical representation of Important features



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.

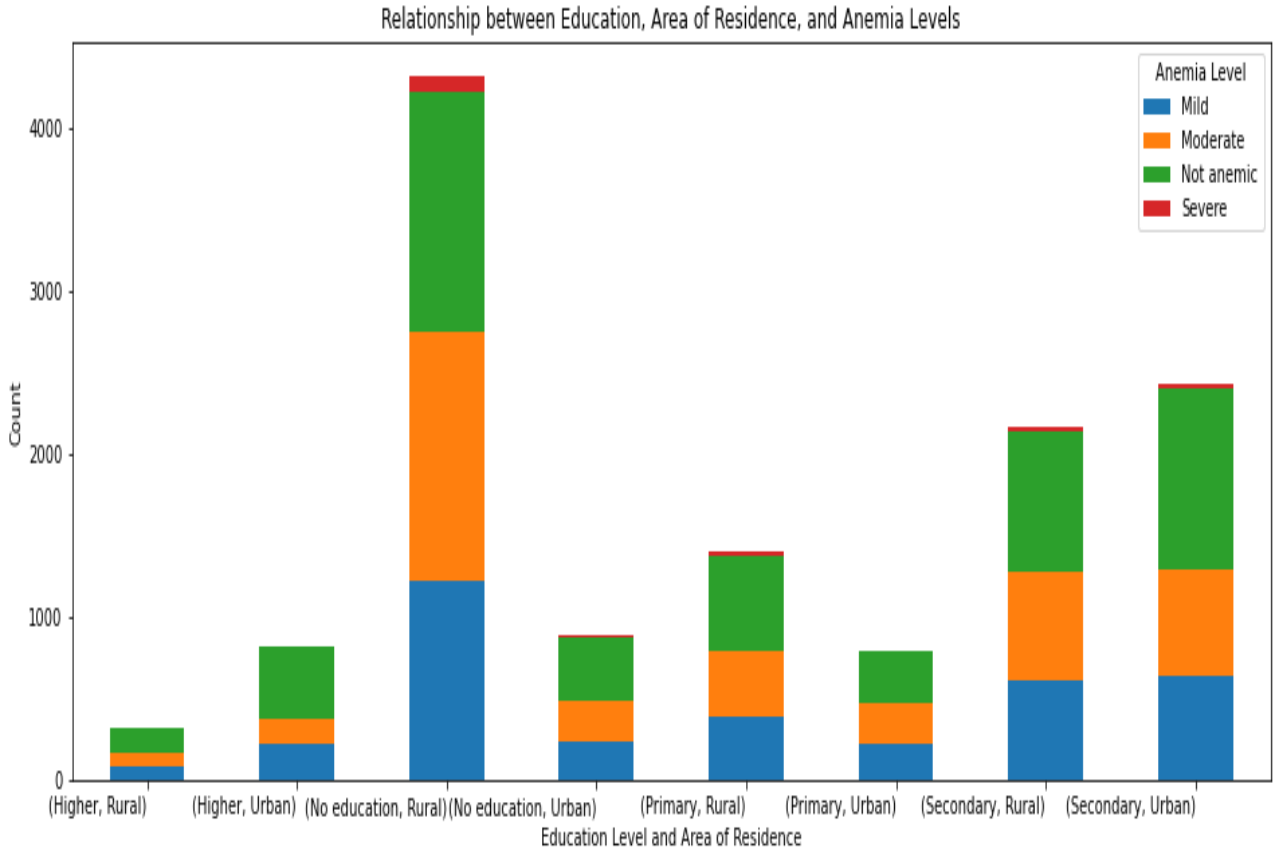
Graphical representation of Important features



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.

Graphical representation of Important features



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

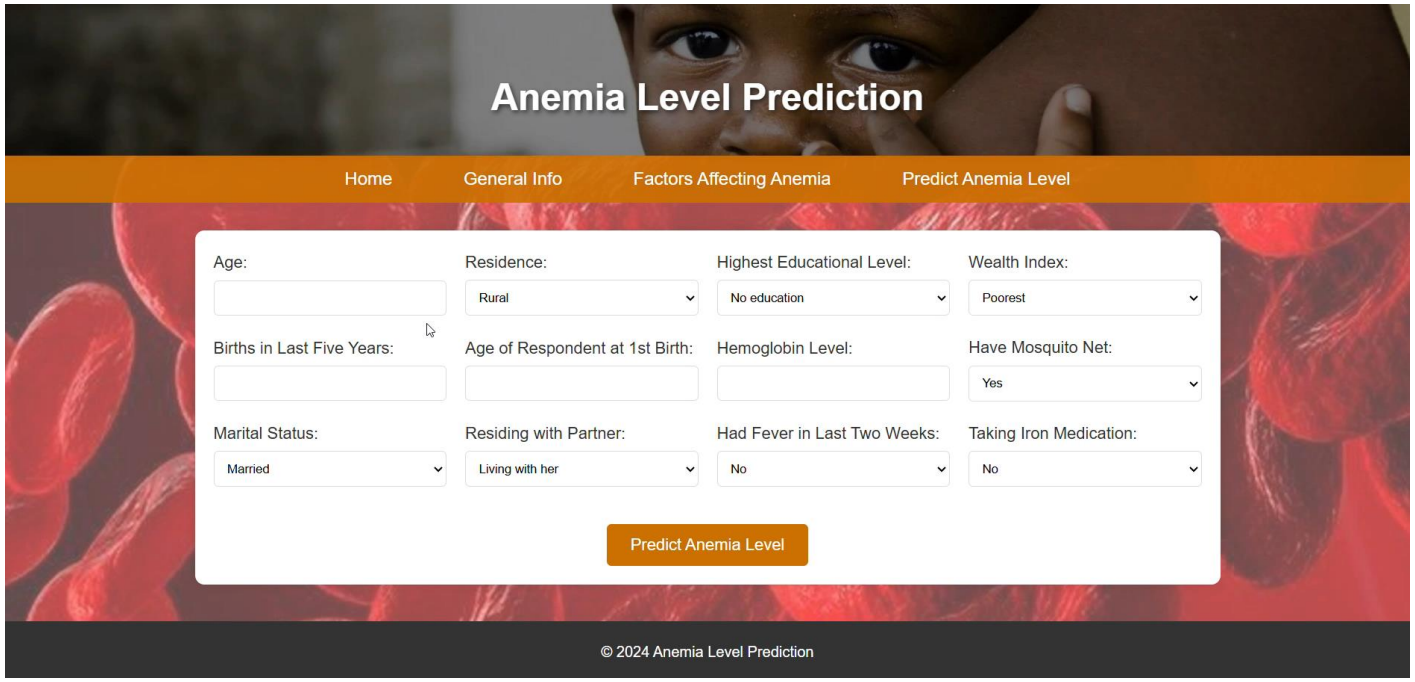
Metrics	Logical reg	Decision tree	Random forest	KNN
Accuracy	0.88	0.93	0.95	0.65
Precision	0.89	0.93	0.95	0.69
Recall	0.88	0.93	0.95	0.67
F1 Score	0.88	0.93	0.95	0.67

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%.

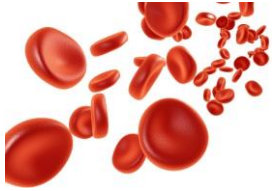
Best Model tunned, improving its accuracy to 96%, making it a strong candidate for deployment

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 that predicts anemia severity among children below 5 years.
- ✓ Analysis of socio-economic determinants reveals that factors such as wealth index and parental education significantly influence anemia severity. These findings guide public health efforts to focus resources on vulnerable demographics, providing valuable insight into the underlying socio-economic drivers of anemia.
- ✓ The model achieved high predictive accuracy, validated through rigorous testing metrics, affirming its readiness for real-world application.
- ✓ The project's findings and the model's application in real-world settings contribute to policy development by offering reliable data that highlight social economic factors associated with anemia.
- ✓ A user-friendly web application was developed as the project's end-user interface. This application allows users to input data, leveraging the pickled model to make predictions and offer tailored recommendations.

Recommendations

- Develop geo-mapping tools to visualize anemia prevalence, enabling precise identification of high-risk regions and resource allocation for targeted interventions.

- Healthcare providers to streamline anemia risk assessment and patient monitoring by integrating predictive models into clinic or hospital dashboards, enhancing decision-making in resource-limited settings.

- Prioritize hemoglobin testing as the primary diagnostic tool through widespread screening programs, particularly for young children, to enable early detection and timely treatment of anemia

- Analyzing socio-economic factors influencing anemia severity highlights the need for targeted intervention strategies. We recommend combining nutritional supplementation with community collaboration to provide subsidized or free iron-rich foods to high-risk households.

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