Model Development Phase

Date	7 August2025
Skill wallet ID	SWUID20250185217
Project Title	Anemia Sense: Leveraging Machine Learning for Precise Anemia Recognition
Maximum Marks	5 Marks

Feature Selection Report:

In the forthcoming update, each dataset feature will be accompanied by a concise yet informative description outlining its relevance, statistical significance, and potential impact on the predictive model. Users will be able to explicitly indicate whether each feature is selected or excluded, along with a clear justification for their decision. This structured approach will not only streamline the feature selection process but also enhance transparency, traceability, and collaboration among stakeholders, ensuring that model inputs are both purposeful and data-driven.

Feature	Description	Selected(Yes/No)	Reasoning
Patient_ID	Unique identifier for each patient	No	For predicting anemia, a patient ID is not required as it does not contribute to the diagnosis
Gender	Patient's gender	Yes	Relevant since gender can influence hemoglobin levels and anemia prevalence.
Hemoglobin	Hemoglobin concentration in g/dL	Yes	A key clinical indicator; low values are strongly associated with anemia.

MCH	Mean Corpuscular Hemoglobin – average hemoglobin per red blood cell	Yes	Useful for distinguishing types of anemia and detecting deficiencies.
MCHC	Mean Corpuscular Hemoglobin Concentration – hemoglobin concentration in red blood cells	Yes	Important for identifying abnormal cell coloration patterns linked to anemia.
MCV	Mean Corpuscular Volume–average size of red blood cells	Yes	Helps classify anemia as microcytic, normocytic, or macrocytic.
Result	Anemia diagnosis result (0 = No Anemia, 1 = Anemia)	Yes	The target variable for predictive modeling – essential for achieving the project's goal.