Feature Selection Report

Date	03 August 2025
Project Title:	Anemia Sense – Machine Learning for Precise Anemia Recognition
Maximum Marks:	5 Marks

Feature Selection Report

In the forthcoming update, each feature is accompanied by a brief description. Selection is indicated by Yes/No, with reasoning provided for transparency in the feature selection process.

Feature	Description	Selected (Yes/No)	Reasoning
Age	Patient age in years	Yes	Age can influence anemia risk; essential for demographic profiling.
Gender	Patient's gender (Male/Female)	Yes	Gender differences affect anemia prevalence and are clinically relevant.
Hemoglobin	Concentration of hemoglobin in blood (g/dL)	Yes	Core biomarker for anemia diagnosis; directly reflects anemia status.

MCH	Mean Corpuscular Hemoglobin – average hemoglobin content per red blood cell	Yes	Useful index for identifying anemia type.
MCHC	Mean Corpuscular Hemoglobin Concentration – hemoglobin concentration in RBCs	Yes	Differentiates anemia categories (e.g., hypochromic vs normochromic anemia).
MCV	Mean Corpuscular Volume – average volume of red blood cells	Yes	Facilitates sub-classification (microcytic/normocytic/macrocytic anemias).
RBC Count	Red Blood Cell count (×10^6/µL)	Yes	Important hematological index—can help refine predictive accuracy for anemia.
Stage	Disease stage (if available in dataset)	No/Optional	Often unavailable or not standardized; may increase noise if inconsistently recorded.
Result	Anemia diagnosis (0: Non-anemic, 1: Anemic)	Yes	**Target variable** for prediction—central to the project objective.
Patient ID	Unique record identifier	No	Not informative for modeling; included only for data tracking.