

## Model Development Phase Template

Date	4 <sup>th</sup> June 2024
Team ID	SWTID1720175375
Project Title	Prediction And Analysis Of Liver Patient Data Using ML
Maximum Marks	5 Marks

### Feature Selection Report

Feature	Description	Selected (Yes/No)	Reasoning
F-score	The harmonic mean of precision and recall. It provides a balanced measure of a model's performance.	Yes	It helps optimize the decision threshold to strike the best balance between correctly identifying diseased individuals and minimizing false positives and negatives.
Total_Proteins & Albumin, <b>Alkaline Phosphatase, Total_Bilirubin, gender</b>	From the jointplots and scatterplots, we find direct relationship with the feature	Yes	There is some correlation between Albumin_and_Globulin_Ratio and Albumin. But its not as high as Total_Proteins & Albumin

Direct_billirubin	<p>The bilirubin that your liver has previously digested is this one.</p> <p>It passes through the small intestine and liver before being eliminated in your urine.</p> <p>The characteristic yellow color of urine is caused by direct bilirubin.</p>	NO	There is high correlation between direct and total billirubin. So dropped the direct billirubin for better model training
AUC-ROC	The curve plots the true positive rate against the false positive rate at various classification thresholds	Yes	Less sensitive to class imbalance compared to accuracy. Medical applications where the trade-off between sensitivity and specificity must be carefully balanced.