# Input Data Features: Swainson's thrush traits (morphology, behaviour, genetics, etc.) Outcome: survival



Remove features with >25% missing

Drop one feature from correlated pairs ( $|r| \ge 0.7$ )

# **Feature Engineering**

Normalize features

## **Pipeline**

KNN Imputation – fill missing values

Train/Test Split (stratified)

SMOTE – balance the dataset

Balanced Random Forest (via RandomizedSearchCV)

### **Model Evaluation**

Classification metrics

Confusion matrix

Threshold tuning for F1 macro

### **Model Interpretation**

Feature importance plot

SHAP dependence plots for key features