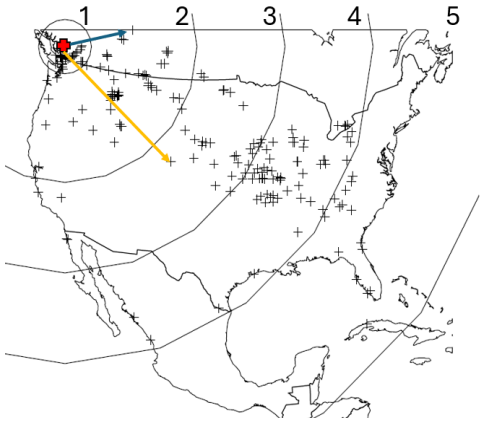
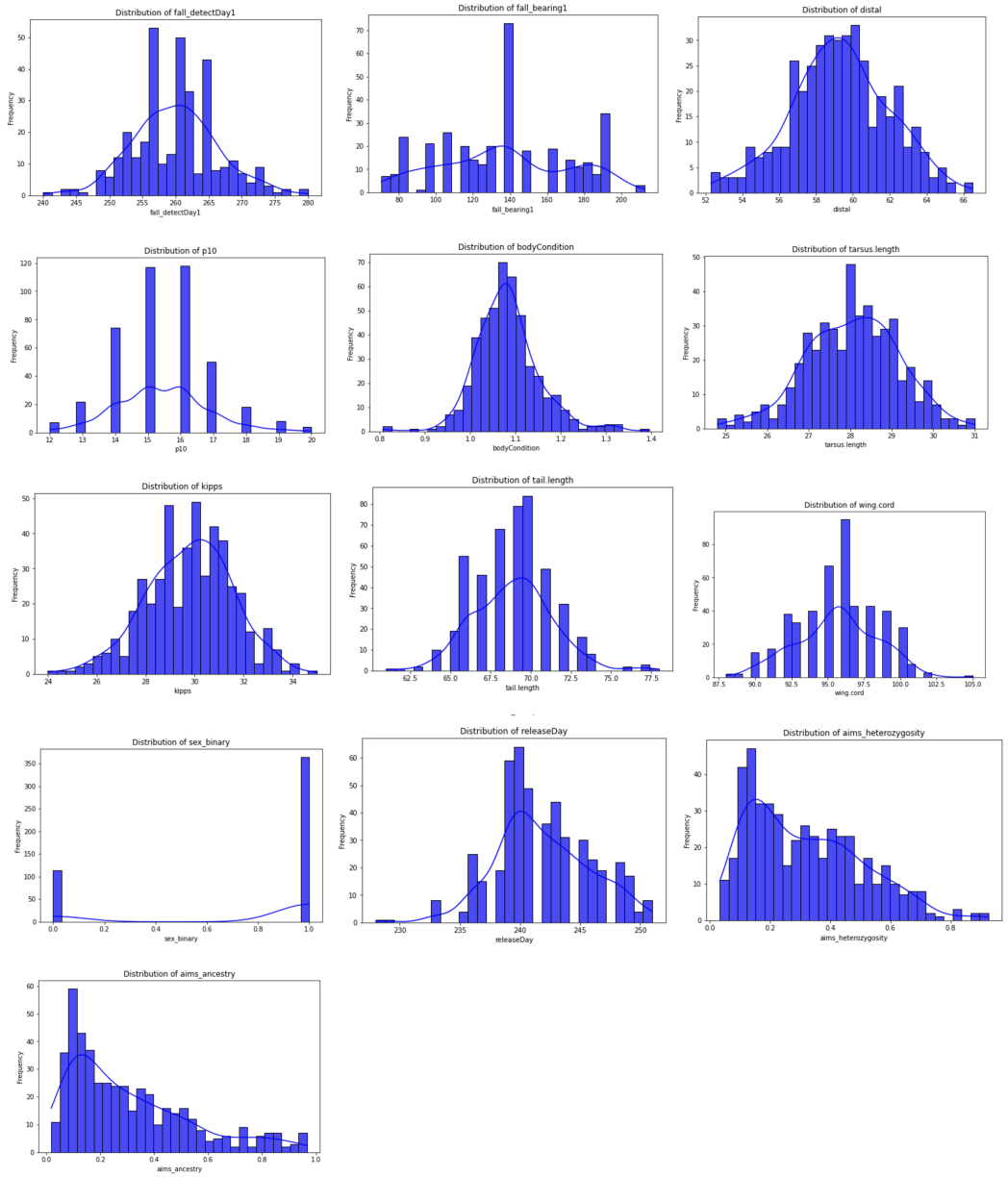
Figure S2: Bearings



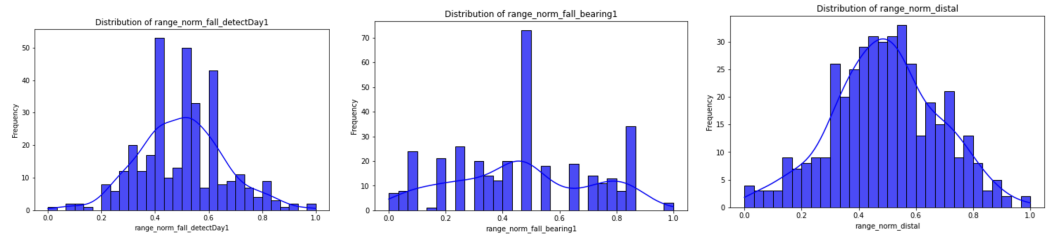
Red cross marks the release site; black crosses indicate bird locations; blue line demonstrates the calculation of bearing 2, and yellow line demonstrates bearing 3.

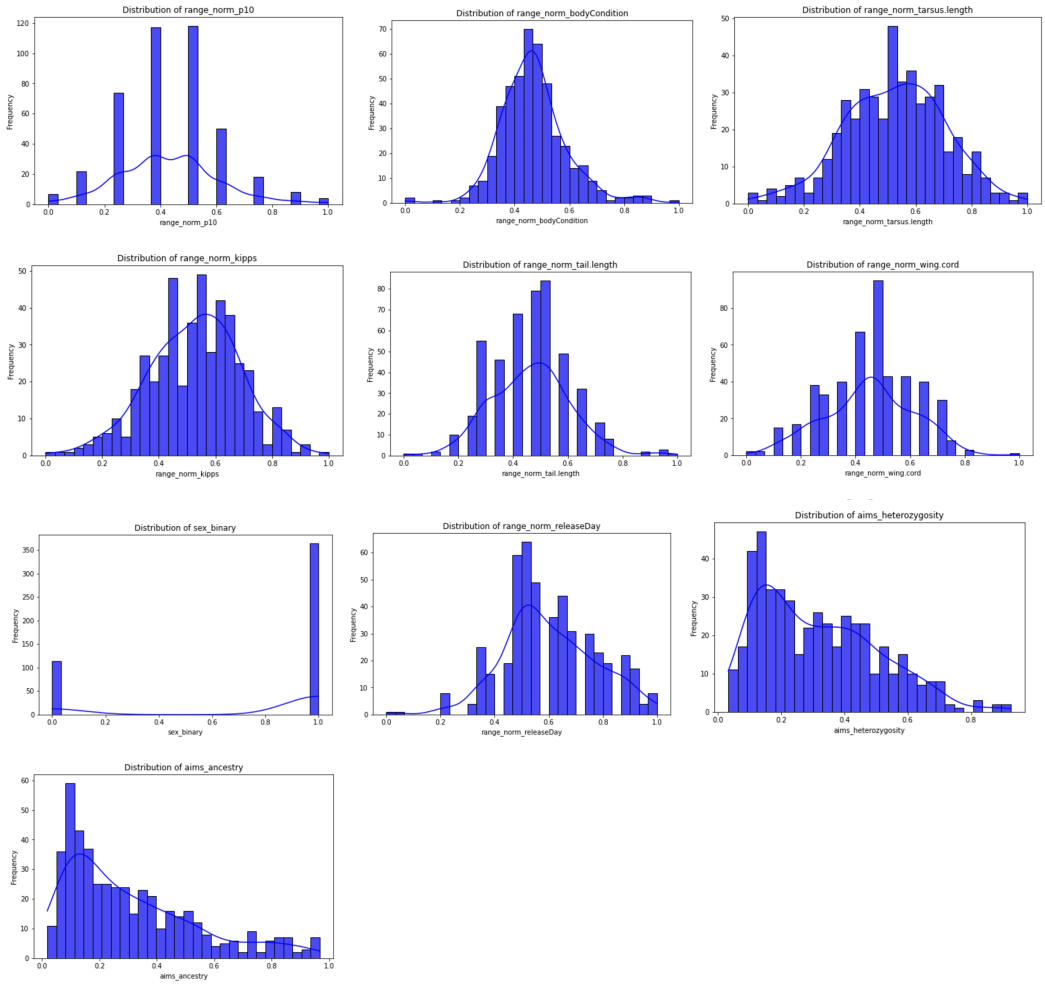
Figure S3: Distributions of key predictive features used in the analysis.

(A)



(B)

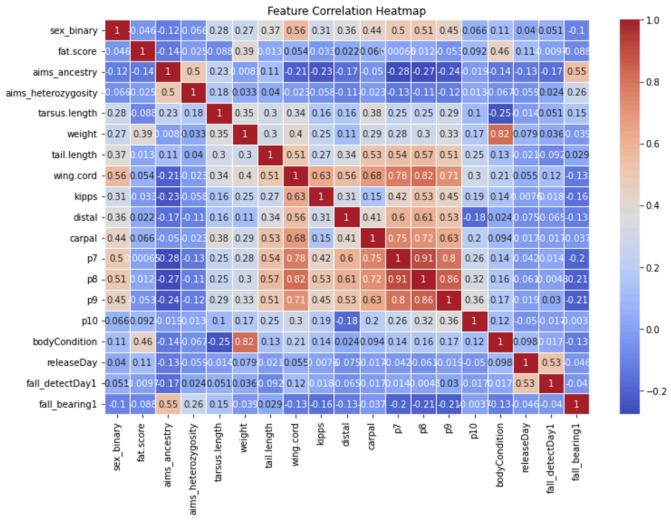




The histograms illustrate the frequency of values for each feature used in the random forest model (A) and the neural network model (B), with kernel density estimates (KDE) overlaid where appropriate to show the underlying probability density.

Figure S4

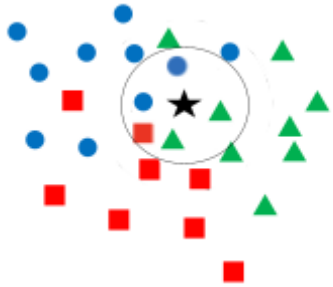
(A)



(B)

Column	Missing Percentage
carpal	37.995825
fall_detectDay1	27.766180
fall_bearing1	27.766180
fat.score	26.096033
distal	13.152401
p10	12.734864
bodyCondition	2.922756
tarsus.length	1.252610
kipps	1.043841
tail.length	0.626305
wing.cord	0.417537
sex_binary	0.000000
aims_ancestry	0.000000
aims_heterozygosity	0.000000
releaseDay	0.000000

(C)



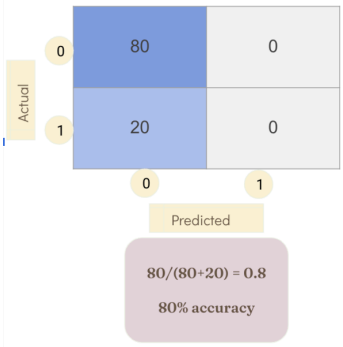
Wing Length	Tail Length	Ancestry
NA	21.3	0.64
32.1	21.2	0.63
32.3	21.4	0.65
31.9	21.1	0.62
32.5	21.5	0.66
32.0	21.3	0.64

$$\text{Imputed Value} = \frac{32.1 + 32.3 + 31.9 + 32.5 + 32.0}{5} = 32.16$$

(A) Feature Correlation. (B) Percentage of missing values for each feature. (C) Visualization of k-Nearest Neighbors (k=5) Imputation Process. The table demonstrates the computation of the missing value (wing length) using the mean of the five nearest neighbors based on their tail length and ancestry values. The imputed value is 32.16, the average of the neighbors' wing length measurements. The accompanying diagram illustrates the five nearest neighbors (colored points) surrounding the missing value (black star), with a black circle indicating the neighborhood boundary for k=5.

Figure S5

(A)



(B)

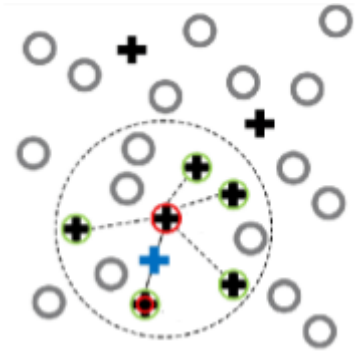


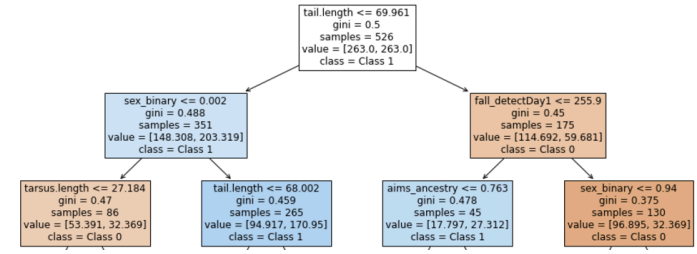
Table (k=5 SMOTE):

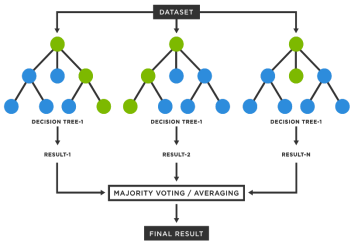
Wing Length	Tail Length	Ancestry	Survival
32.2	21.3	0.64	1
32.1	21.2	0.63	1
32.3	21.4	0.65	1
31.9	21.1	0.62	1
32.5	21.5	0.66	1
32.0	21.3	0.64	1
Synthetic Samples (Examples)			
32.15	21.25	0.635	1

Visualization of k-Nearest Neighbors (k=5) for SMOTE Oversampling: The table demonstrates the generation of synthetic samples by interpolating between a randomly selected minority class sample and its five nearest neighbors. The diagram provides a detailed illustration: gray circles represent majority class samples, black crosses represent minority class samples, and the black cross circled in red is the randomly selected minority class sample. The black crosses with green circles are the five nearest neighbors of the selected minority sample, while the black cross with both a red and green circle indicates the randomly chosen sample from these neighbors. The blue cross denotes the generated synthetic minority sample, placed along the lines connecting the selected minority sample and its chosen neighbor. A dashed black circle indicates the neighborhood boundary for k=5.

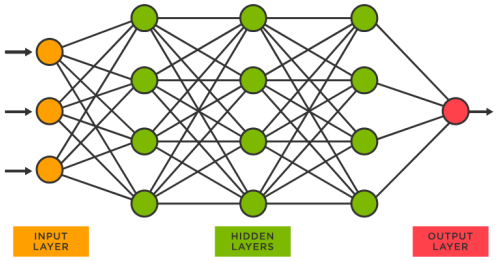
Figure S6

(A) (B)





(C)



(A) Decision Tree Example (B) Illustration of the Random Forest model. (C) Schematic representation of the neural network model.