

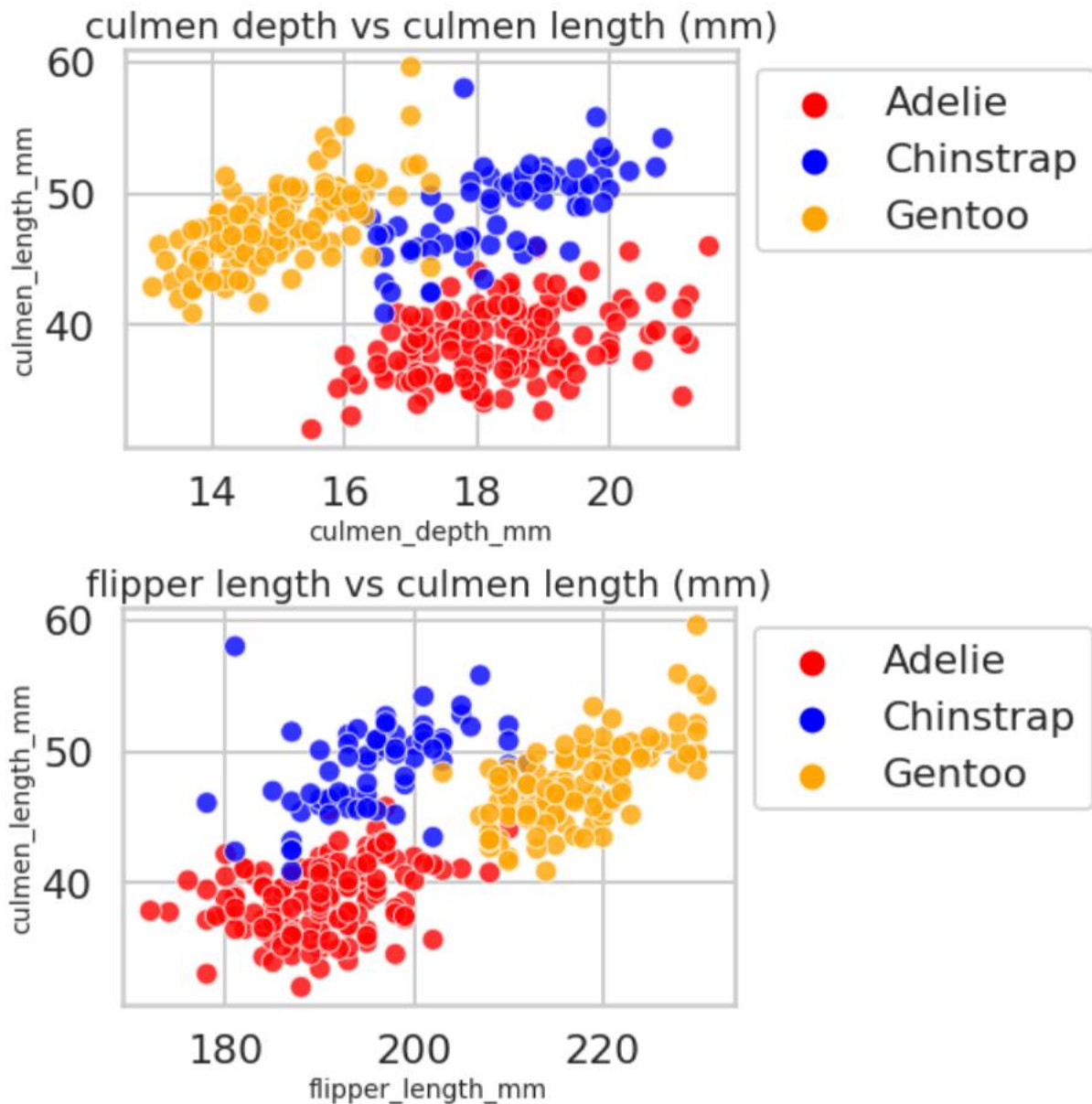
Penguin Species Prediction

Using Feature Analysis, Tree and Linear
Logistic Regression

by **Rajat Chakraborty**

1. Feature Analysis

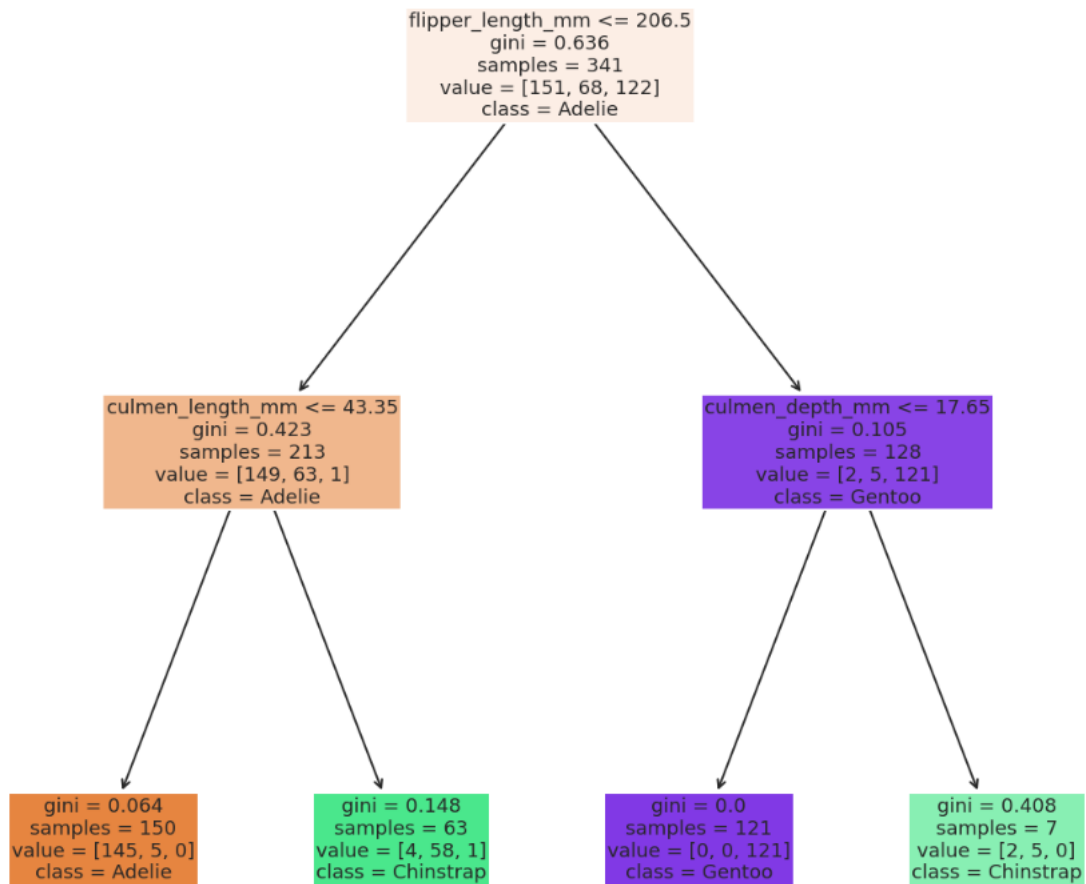
a. Visualization of Features



Flipper length vs culmen length shows small number of overlaps. So, it is good to choose these two features for classification. Even culmen length vs culmen depth also shows low overlaps.

a. Tree Rule

: If flipper length is greater than 206.5 mm and culmen depth is less than 17.65 mm, it is 'Gentoo'.



2. Linear Logistic Regression

5-fold Cross-Validation Accuracy: **99.41%**