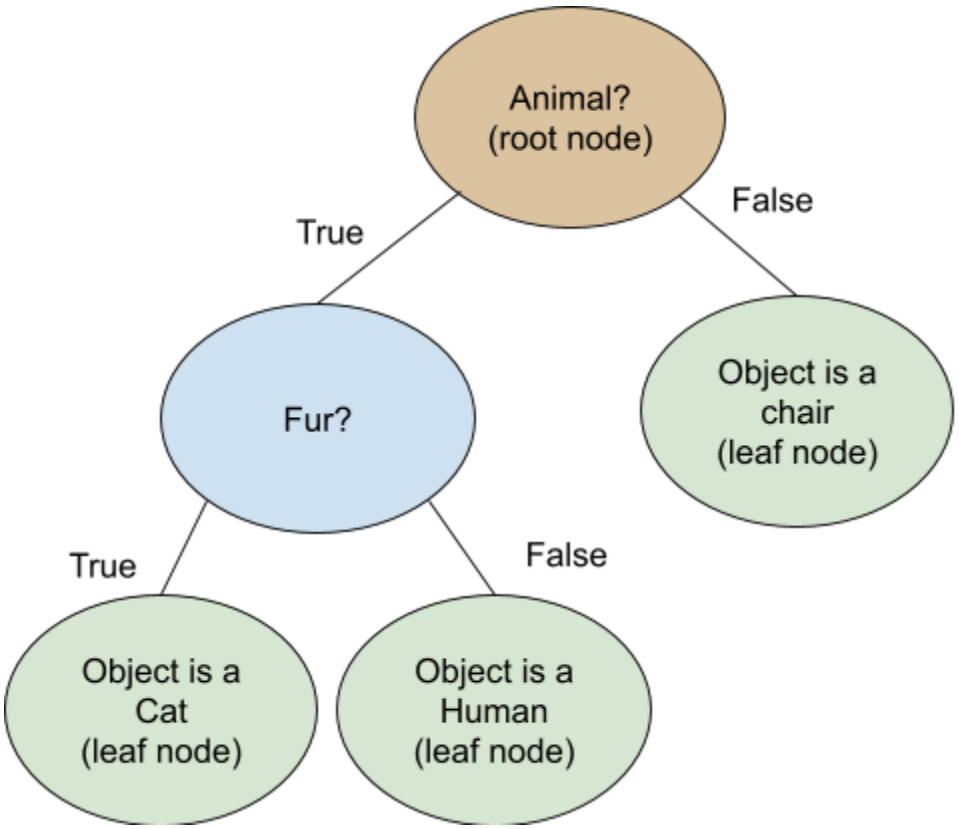
Decision Tree Classification

I'm going second for once \o/

Classification is giving a qualitative tag to an object based on characteristics of that object, like whether an email is spam or ham

Decision tree classification works well with both numerical and categorical data.

Really terrible Decision Tree Classification ↓



The difference between decision tree classification and regression is that regression has a numerical output whereas classification has a classification output.

Naive Bayes

Naive Bayes finds the probability of classifications based on the features.
Featurize: Turning data into numbers that can be interpreted by a computer
vectorize: Turning data into numbers that can be used by vector algorithms

To find the probability of a label occurring in a data set you take the number of instances of the label and divide it by total number of rows in the dataset.
The equation would look like this: $prob = num\ of\ label\ /\ total$
In code with the dataframe named DF the pseudocode would be:

Get the num of label
Get the len of DF
Do the calculation between the variables