Decision Trees Classification
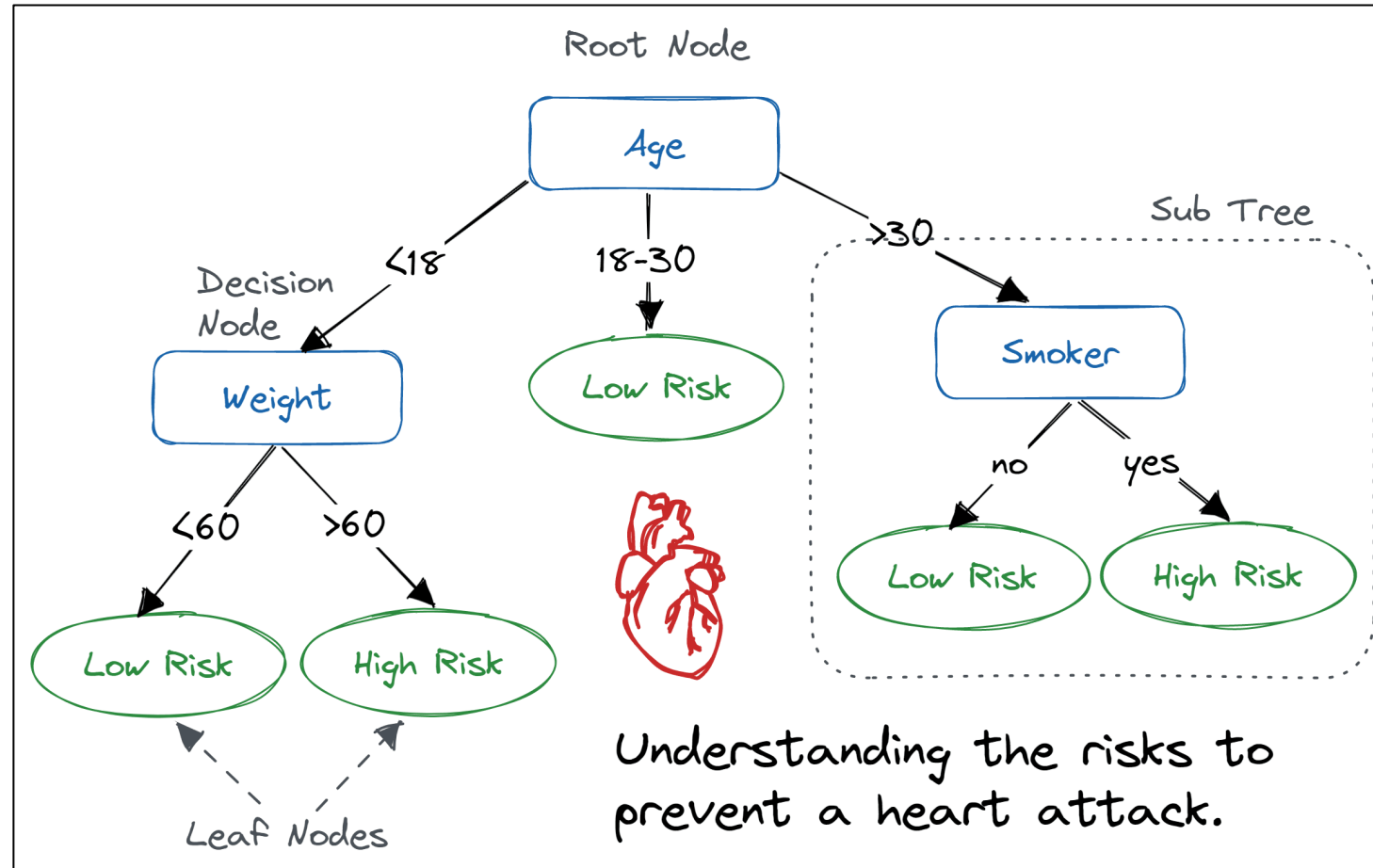
UTKARSH GAIKWAD

Classification vs Regression
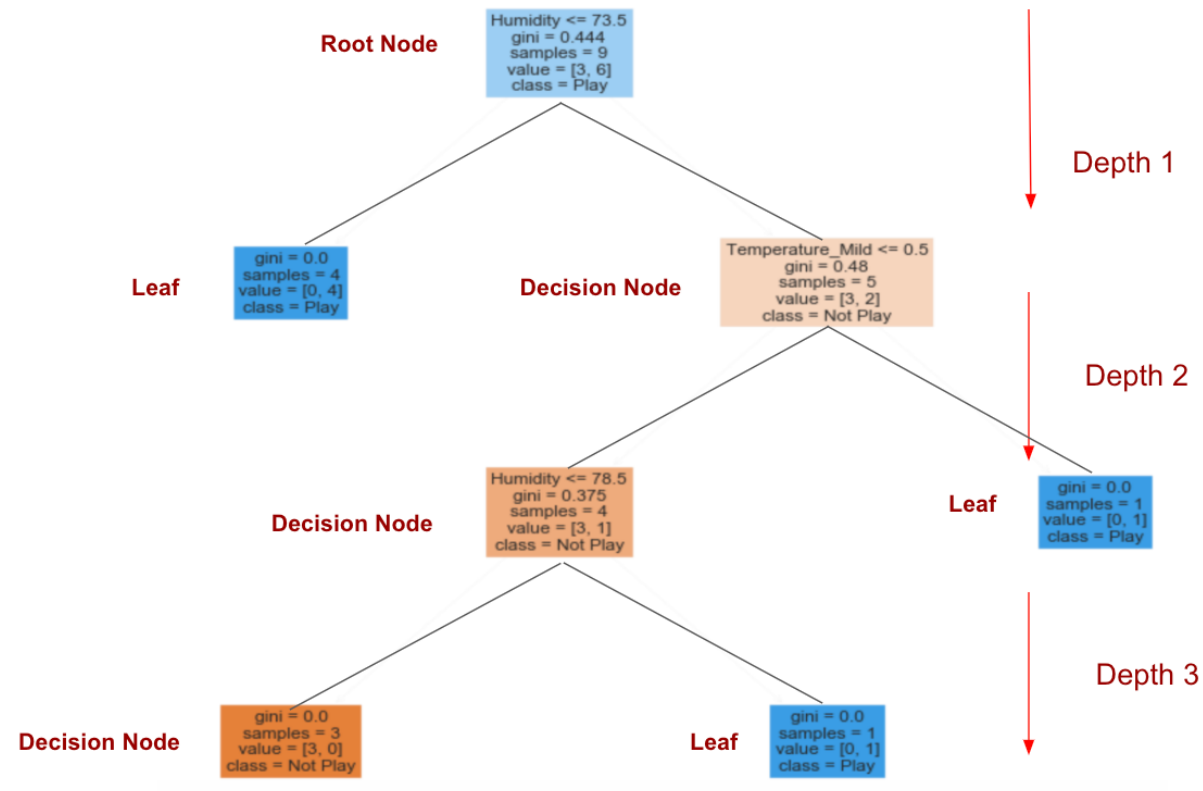
Dependent Feature
Continuous (Numerical Features)

Regression
House price prediction

Dependent Feature
Categorical (Text Features)

Classification

Binary Classification
Eg. Email Spam or not spam

Multiclass Classification
Eg. Separating IRIS Species
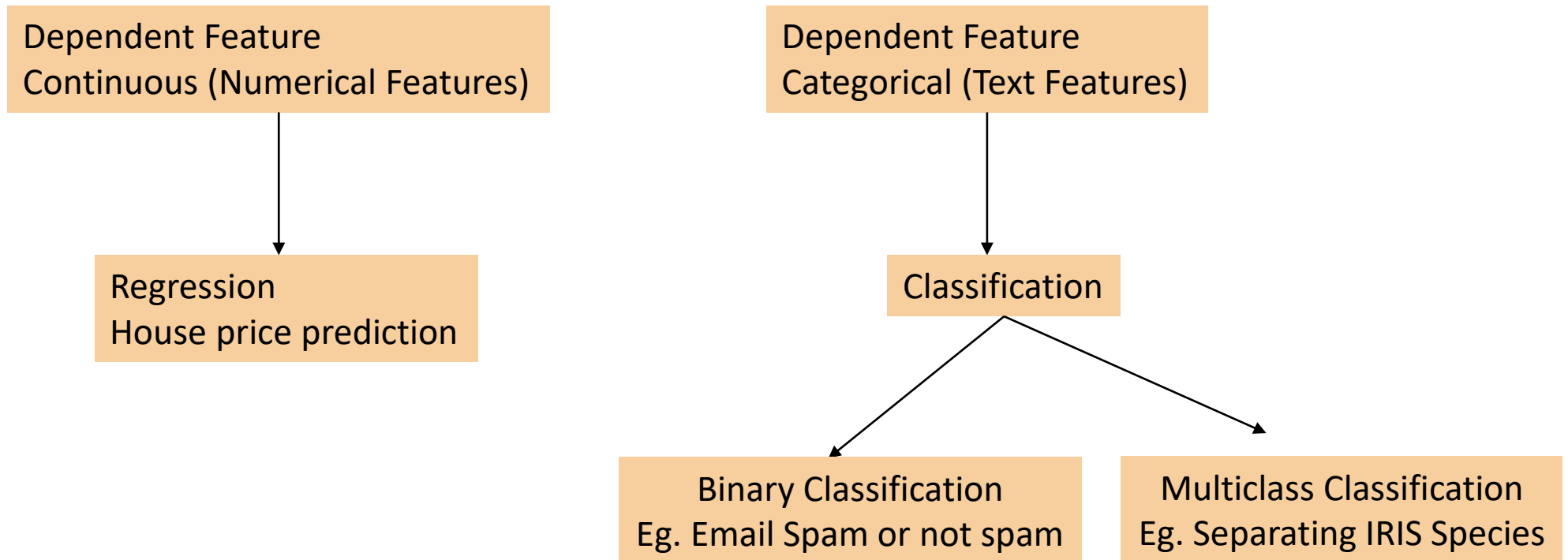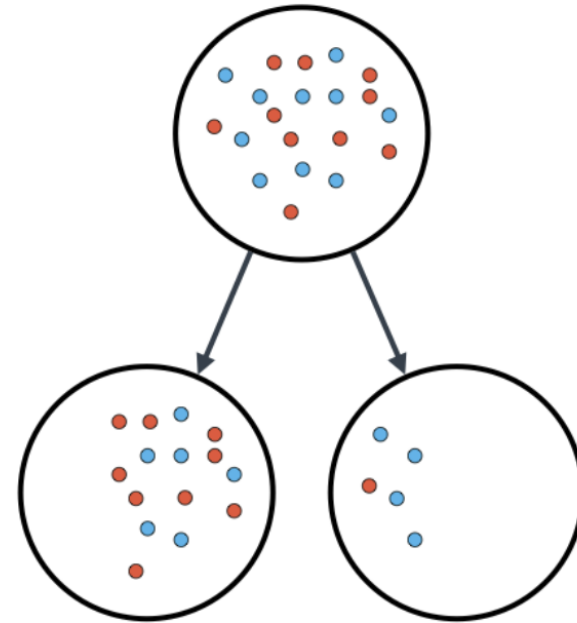
Example of a Classification Tree

Decision Tree Max Depth

Decision Tree Min Samples Split
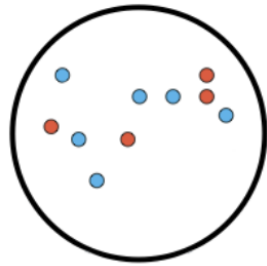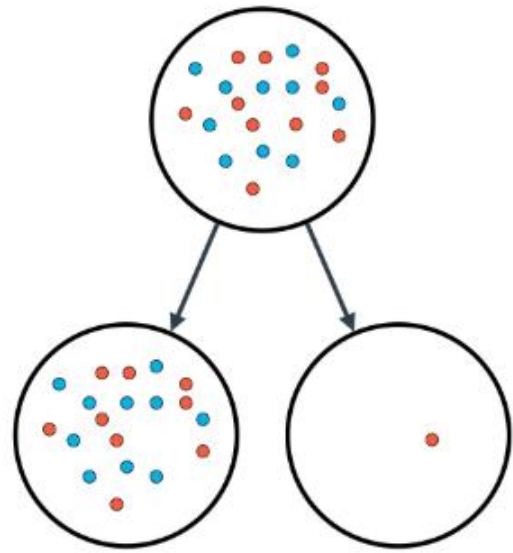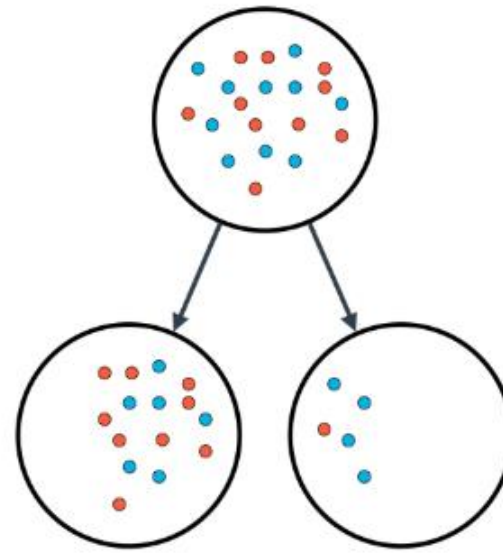

No split!

Minimum number of samples to split = 11   Minimum number of samples to split = 11

Min Samples Leaf

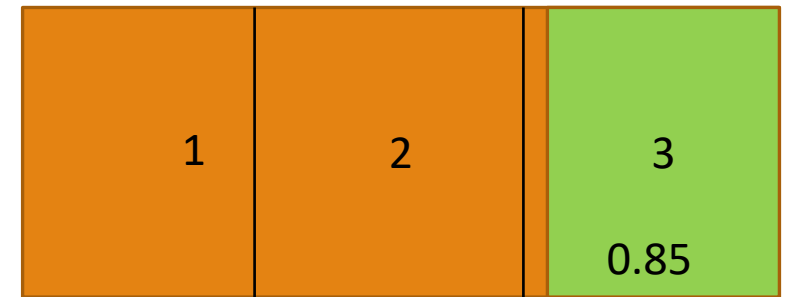
Minimum samples per leaf = 1   Minimum samples per leaf = 5
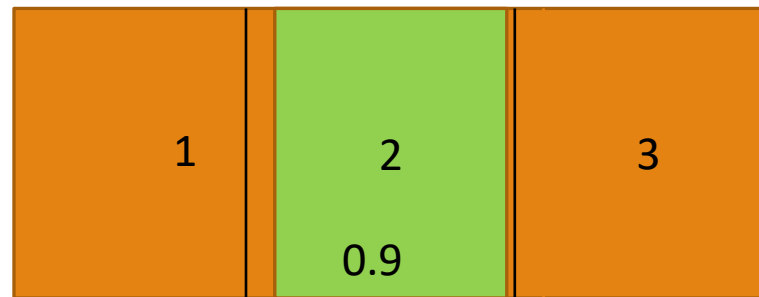
Minimum number of samples per leaf

Gini vs Entropy

$$Gini : Gini(E) = 1 - \sum_{j=1}^{c} p_j^2$$

$$Entropy : H(E) = -\sum_{j=1}^{c} p_j \log p_j$$

K Fold Cross validation

Thank you

PING ME ON SKYPE FOR ANY QUERIES