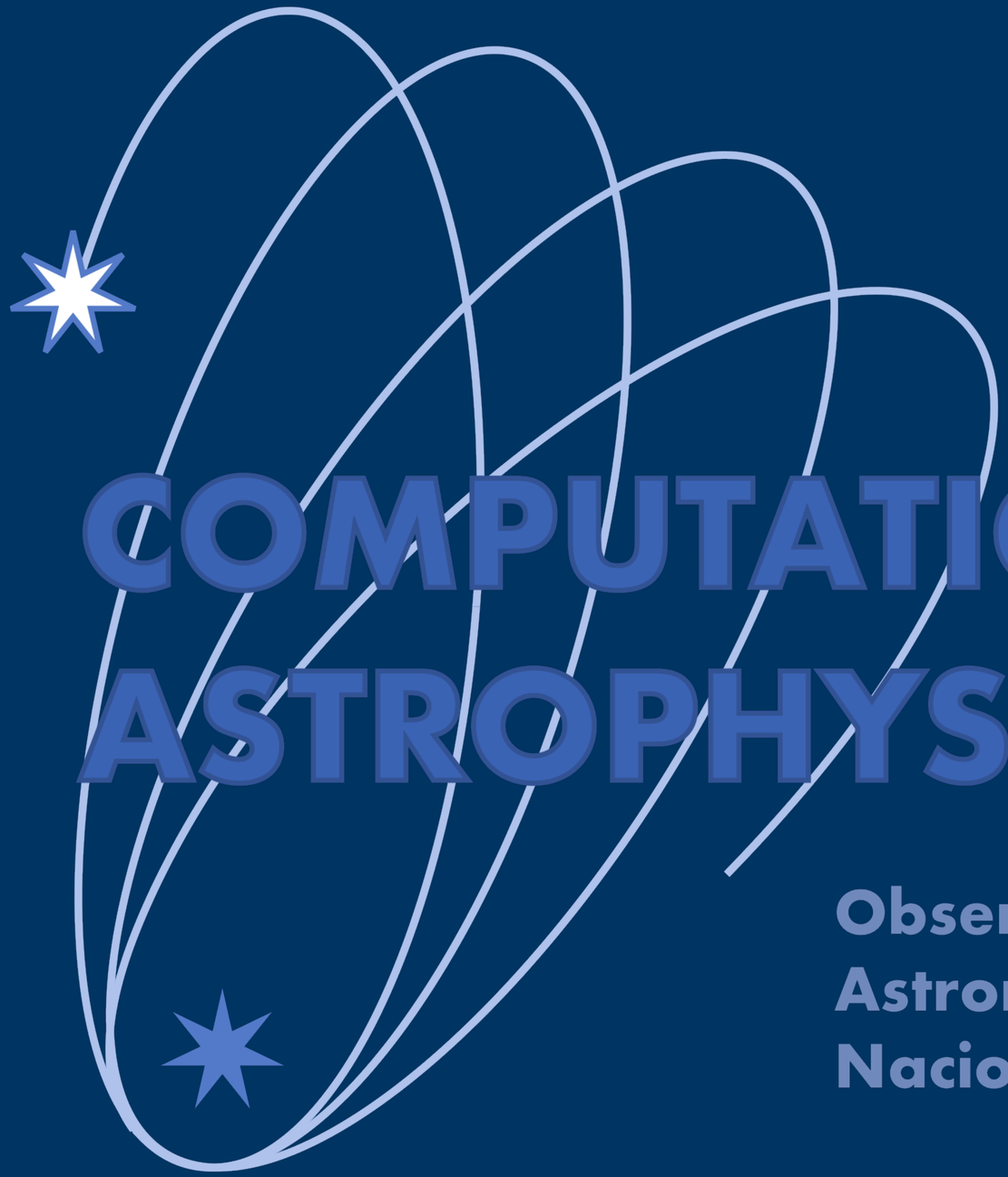


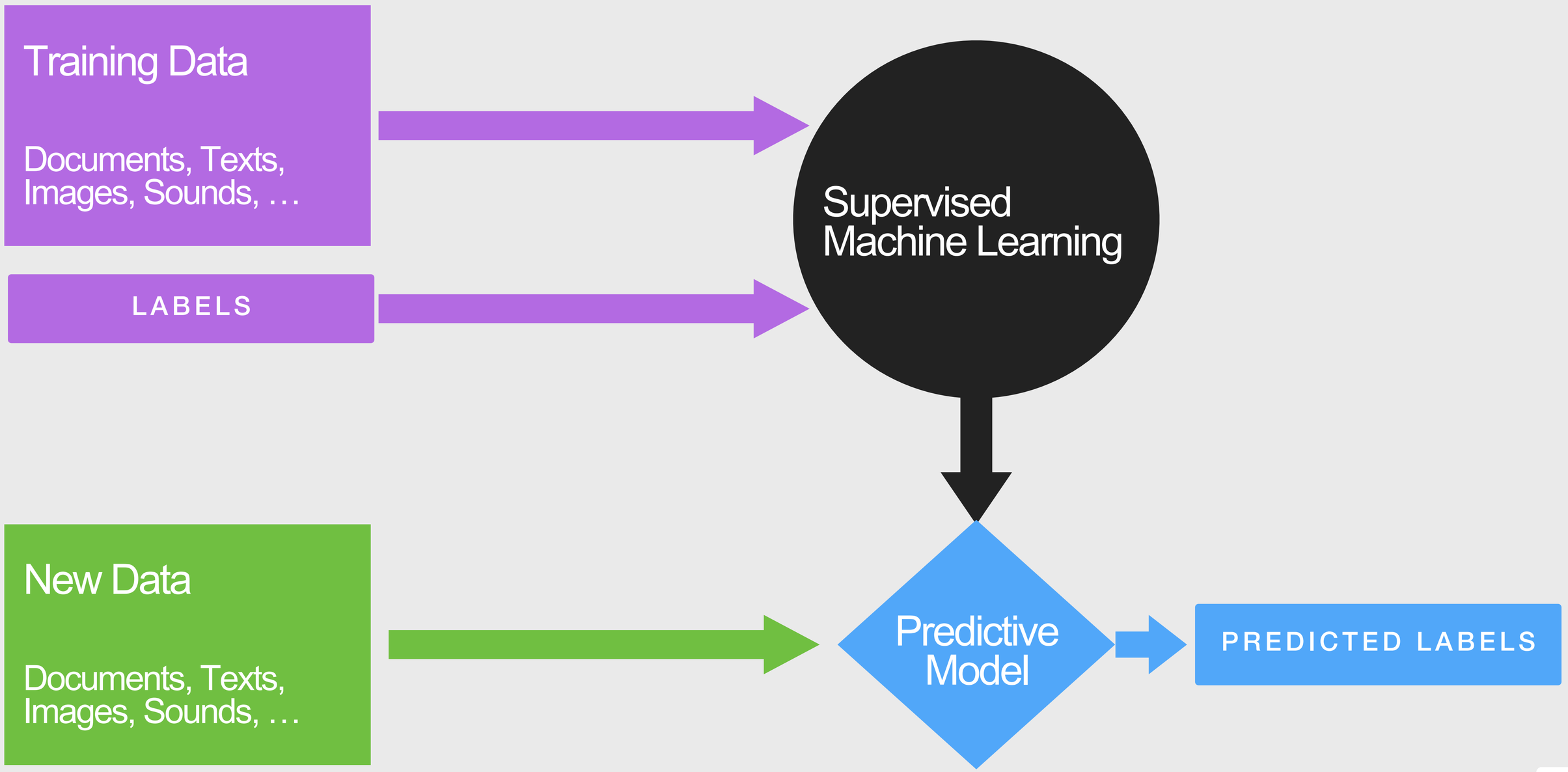
COMPUTATIONAL ASTROPHYSICS

Observatorio
Astronómico
Nacional

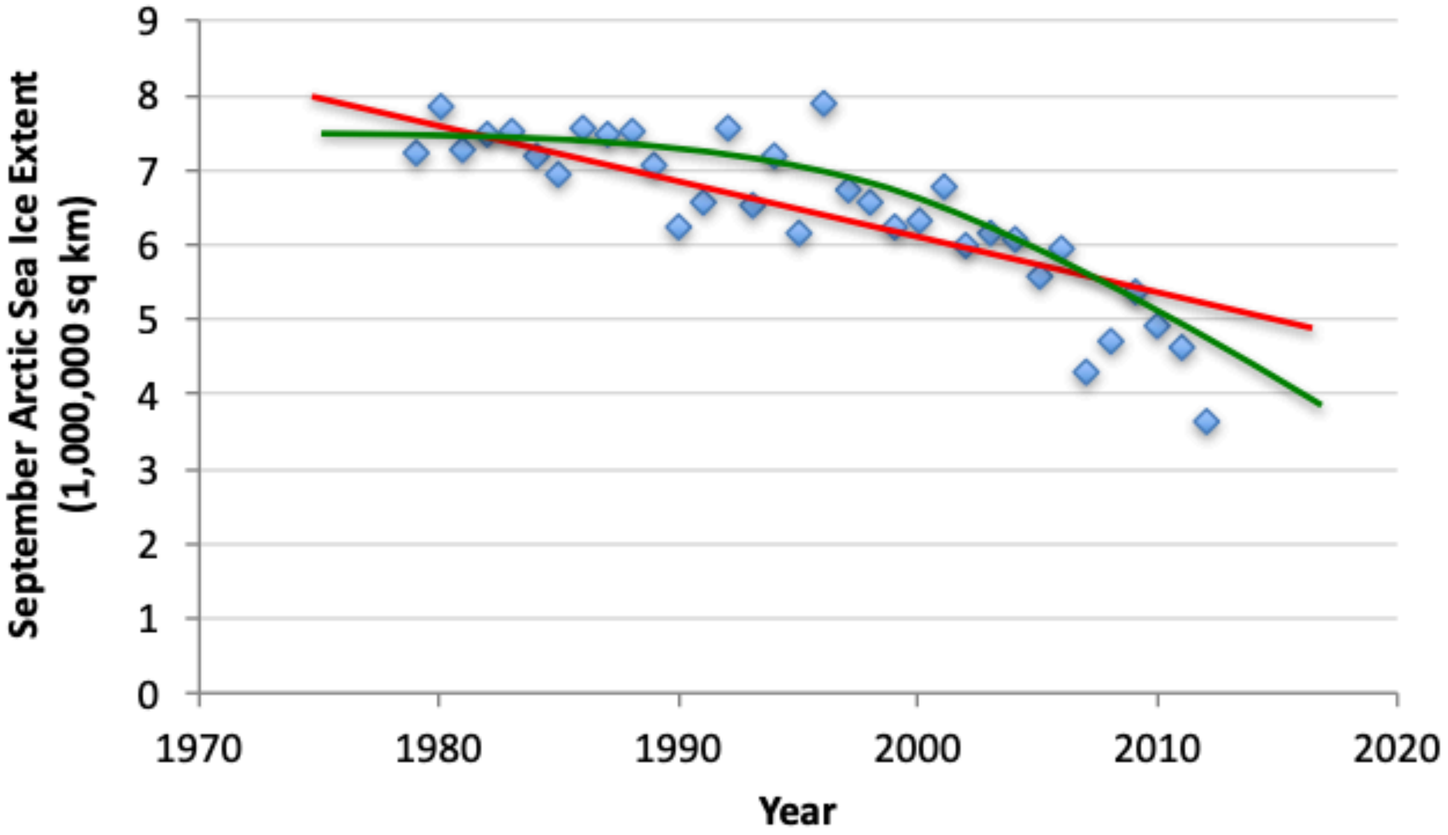
Computational Astrophysics

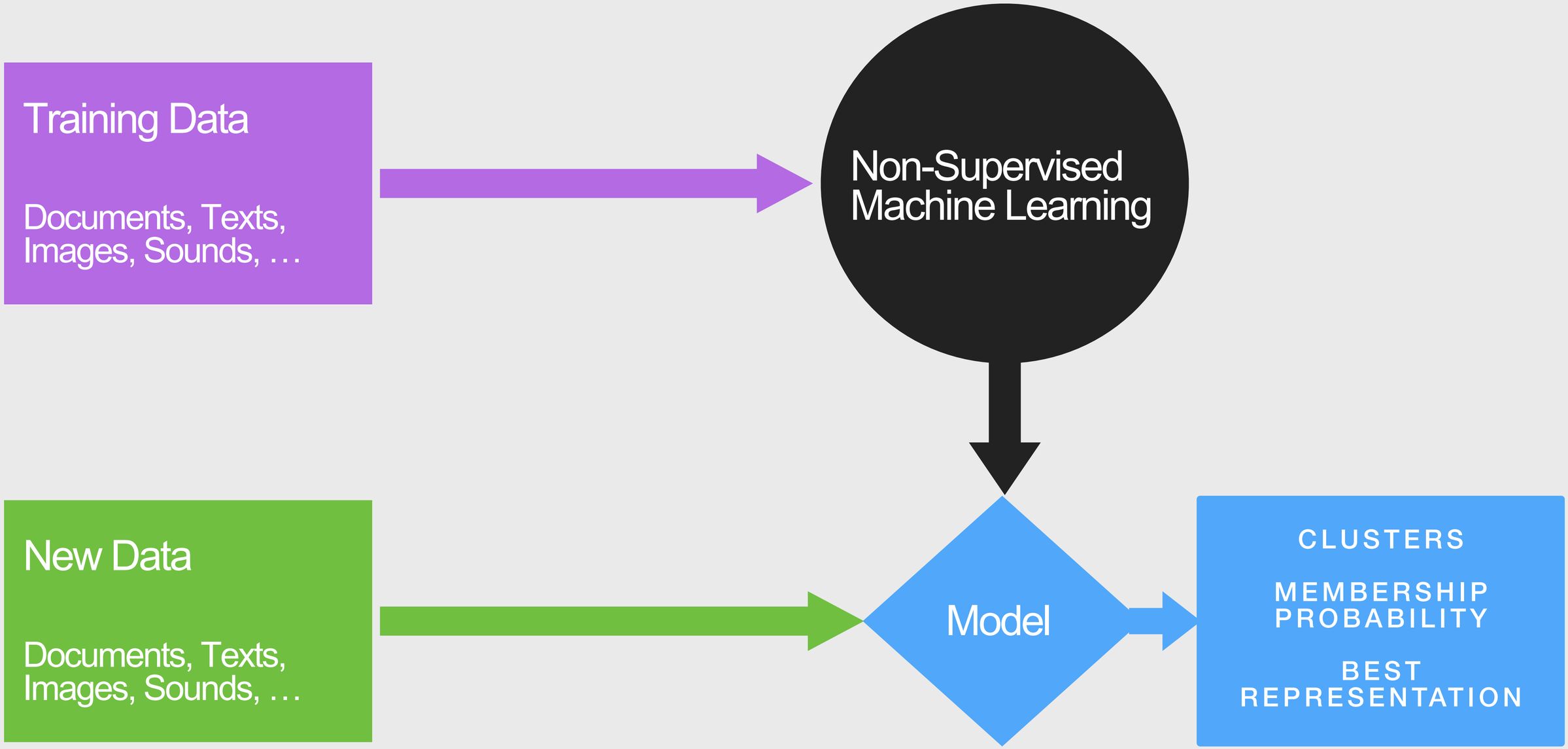
09A. Decision Trees

Eduard Larrañaga
Observatorio Astronómico Nacional
Universidad Nacional de Colombia

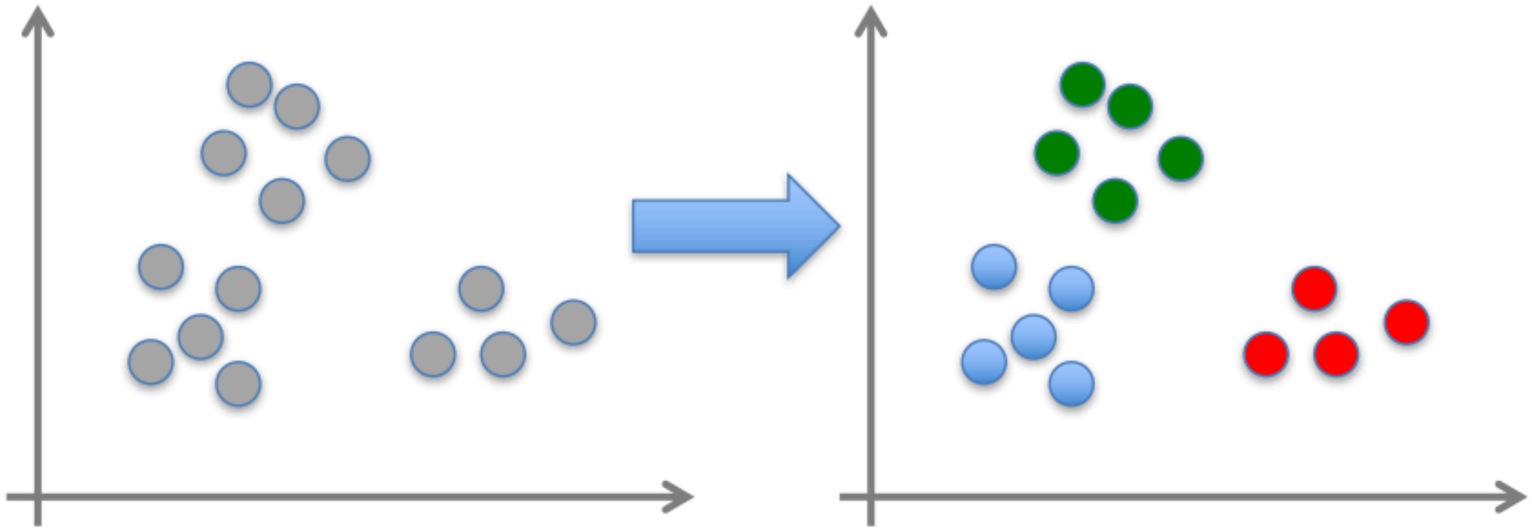

Training Data
Documents, Texts, Images, Sounds, …
LABELS
Supervised Machine Learning
New Data
Documents, Texts, Images, Sounds, …
Predictive Model
PREDICTED LABELS

Regression (Supervised ML)

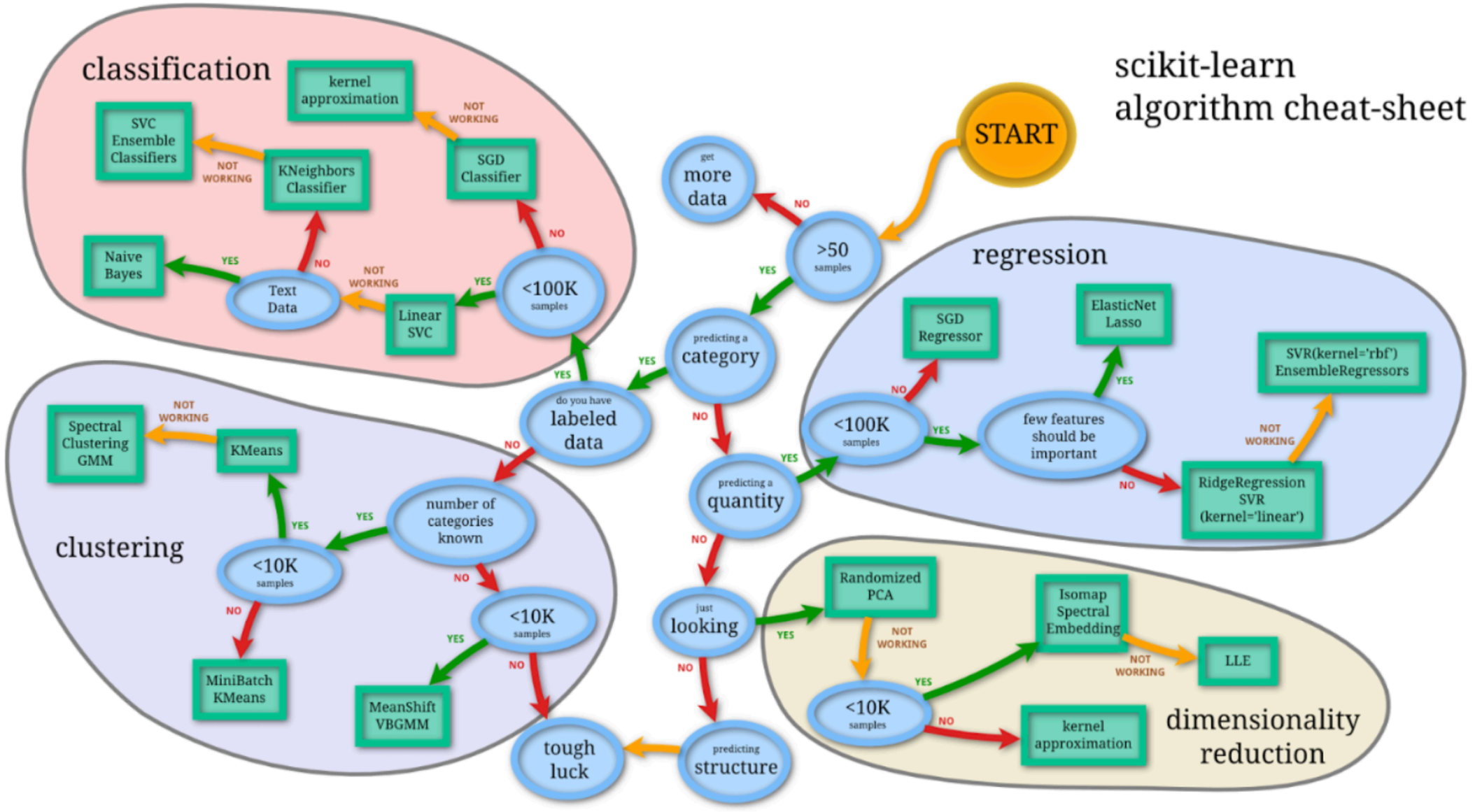



Training Data
Documents, Texts, Images, Sounds, …
Non-Supervised Machine Learning
New Data
Documents, Texts, Images, Sounds, …
Model
CLUSTERS
MEMBERSHIP PROBABILITY
BEST REPRESENTATION

Clustering (No-Supervised ML)

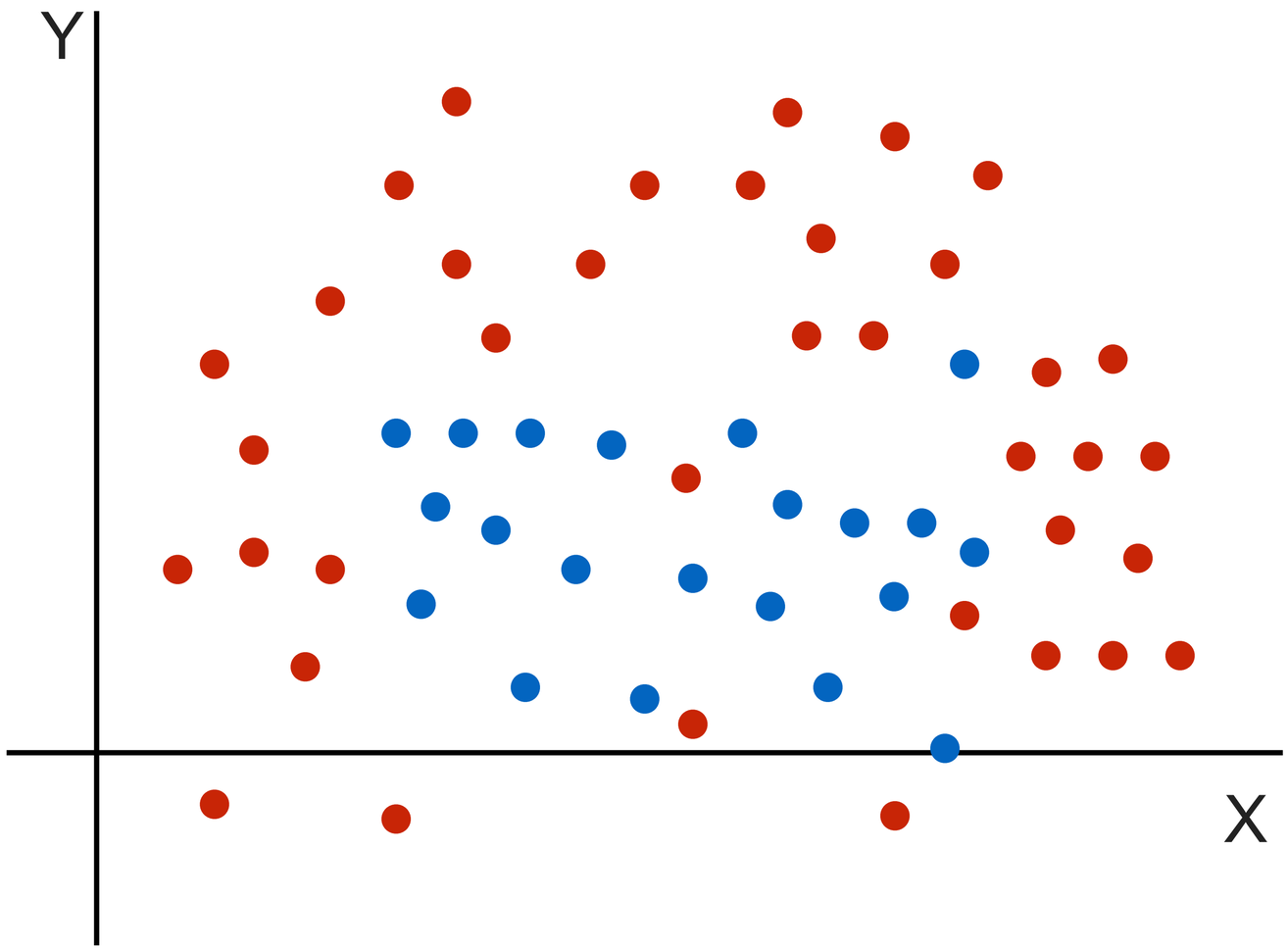


python
scikit
learn

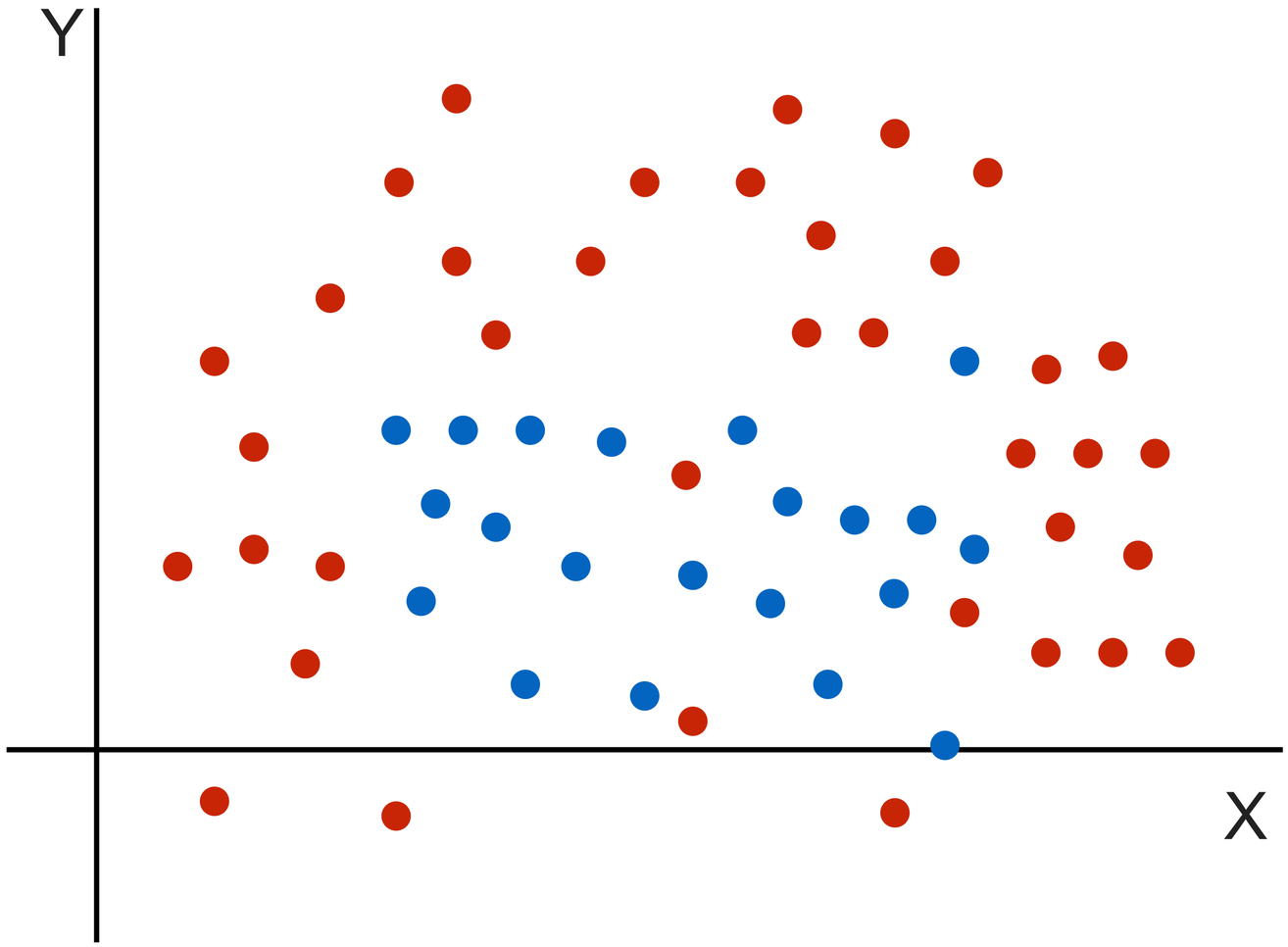

scikit-learn
algorithm cheat-sheet
START
get more data
>50 samples
predicting a category
do you have labeled data
predicting a quantity
just looking
predicting structure
tough luck
classification
kernel approximation
SVC Ensemble Classifiers
KNeighbors Classifier
SGD Classifier
Naive Bayes
Text Data
Linear SVC
<100K samples
clustering
Spectral Clustering GMM
KMeans
number of categories known
<10K samples
MiniBatch KMeans
<10K samples
MeanShift VBGMM
regression
SGD Regressor
ElasticNet Lasso
<100K samples
few features should be important
SVR(kernel='rbf') EnsembleRegressors
RidgeRegression SVR (kernel='linear')
dimensionality reduction
Randomized PCA
Isomap Spectral Embedding
LLE
<10K samples
kernel approximation
YES
NO
NOT WORKING

Decision Trees

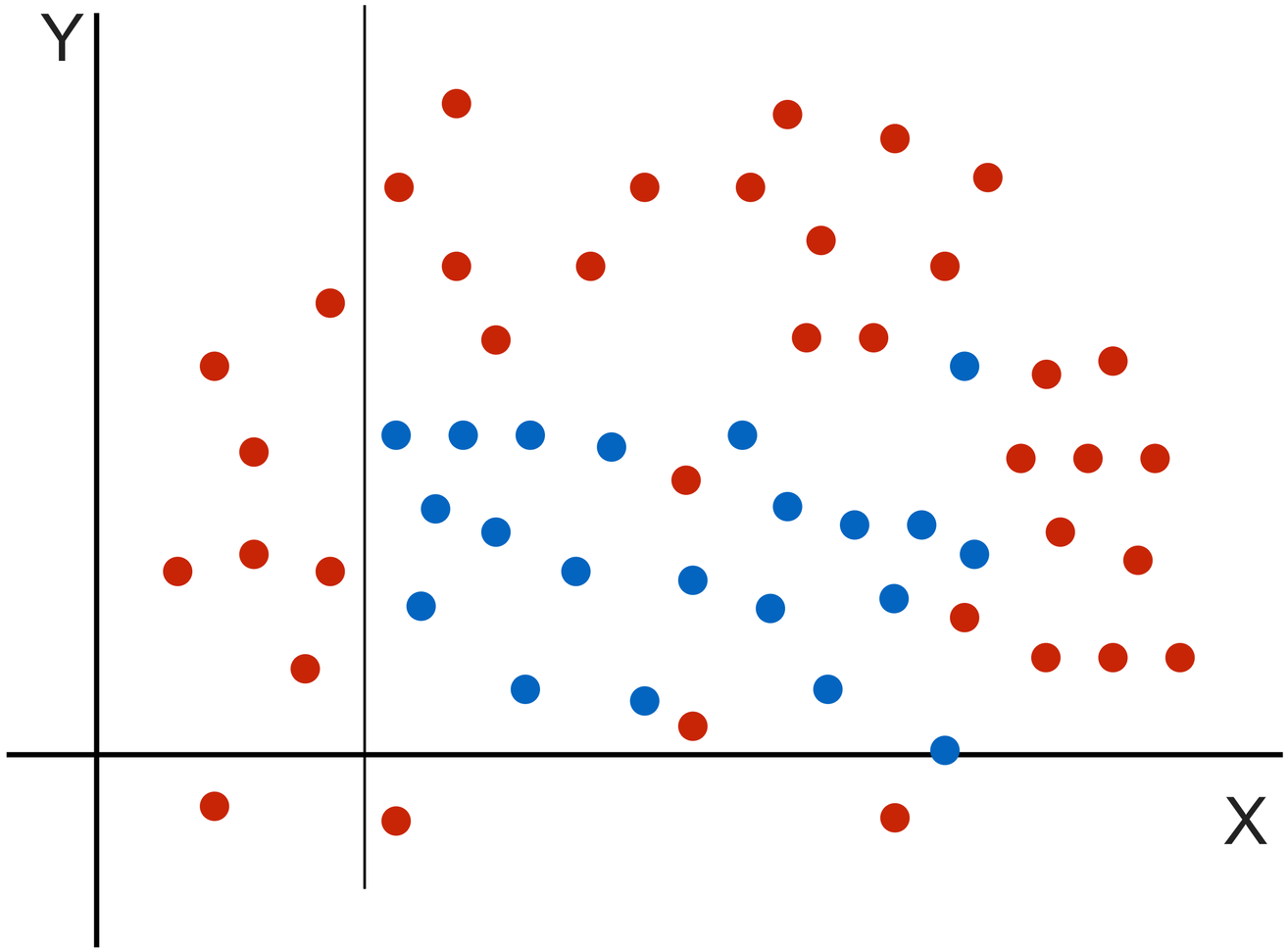
Decision Trees

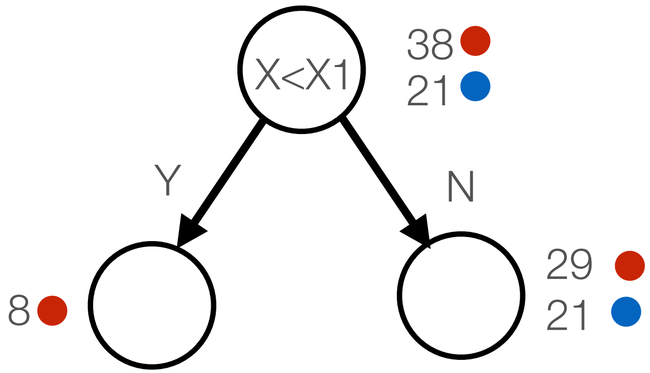


Decision Trees

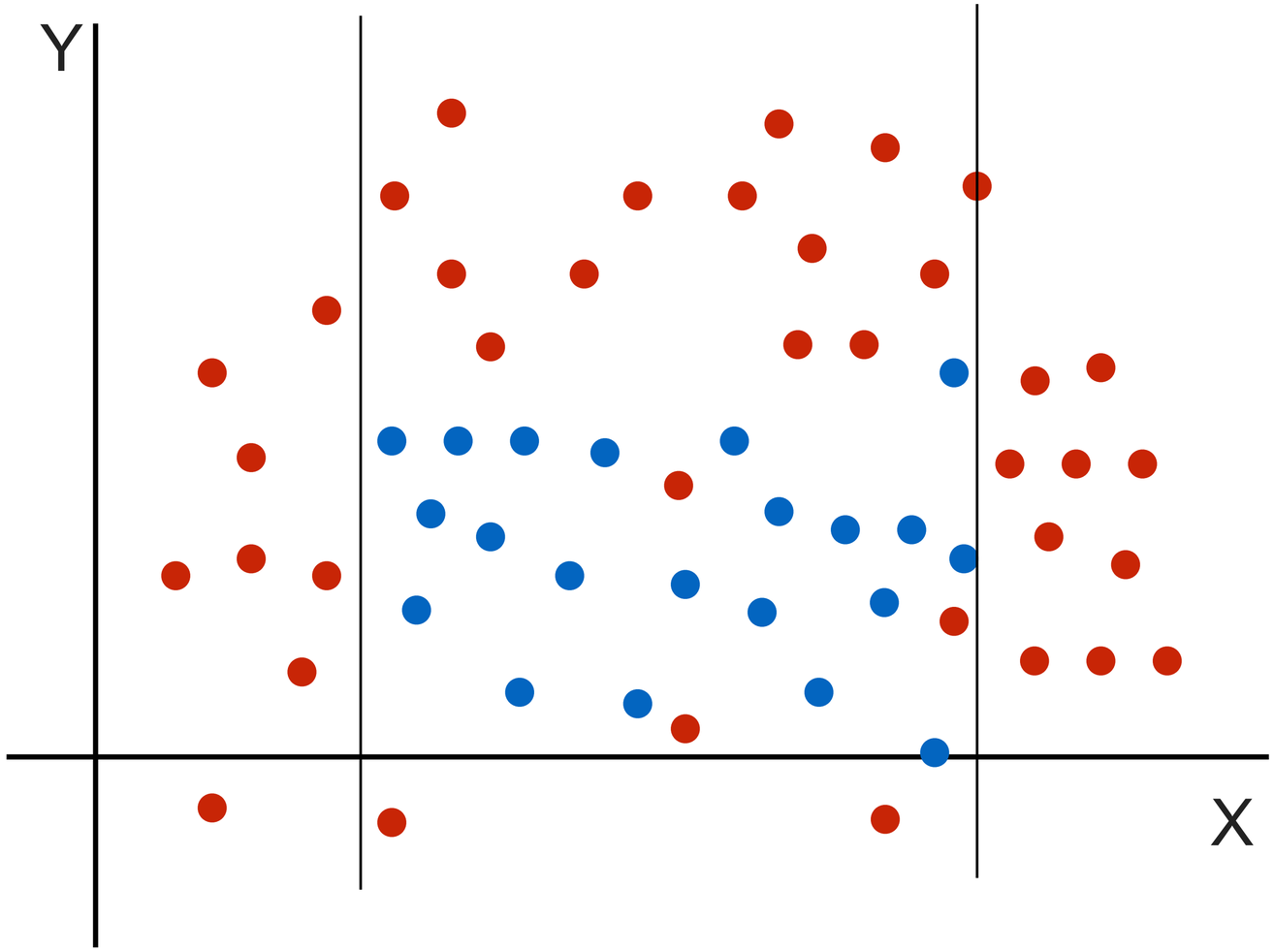


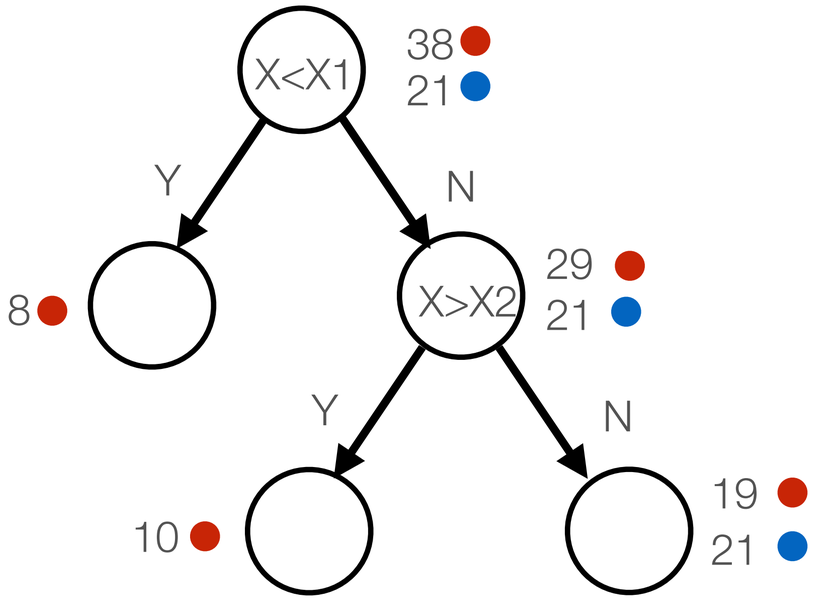
Decision Trees



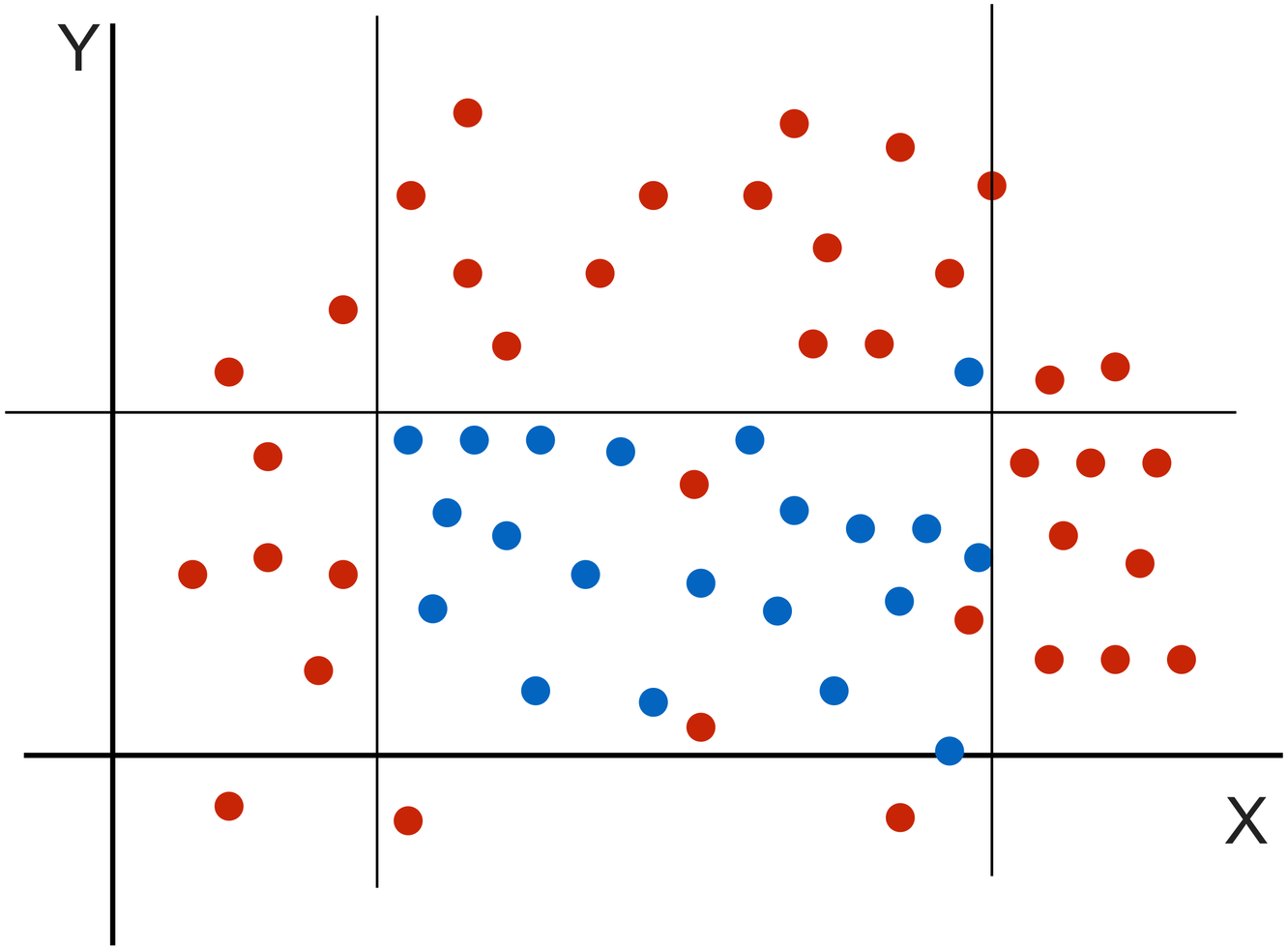


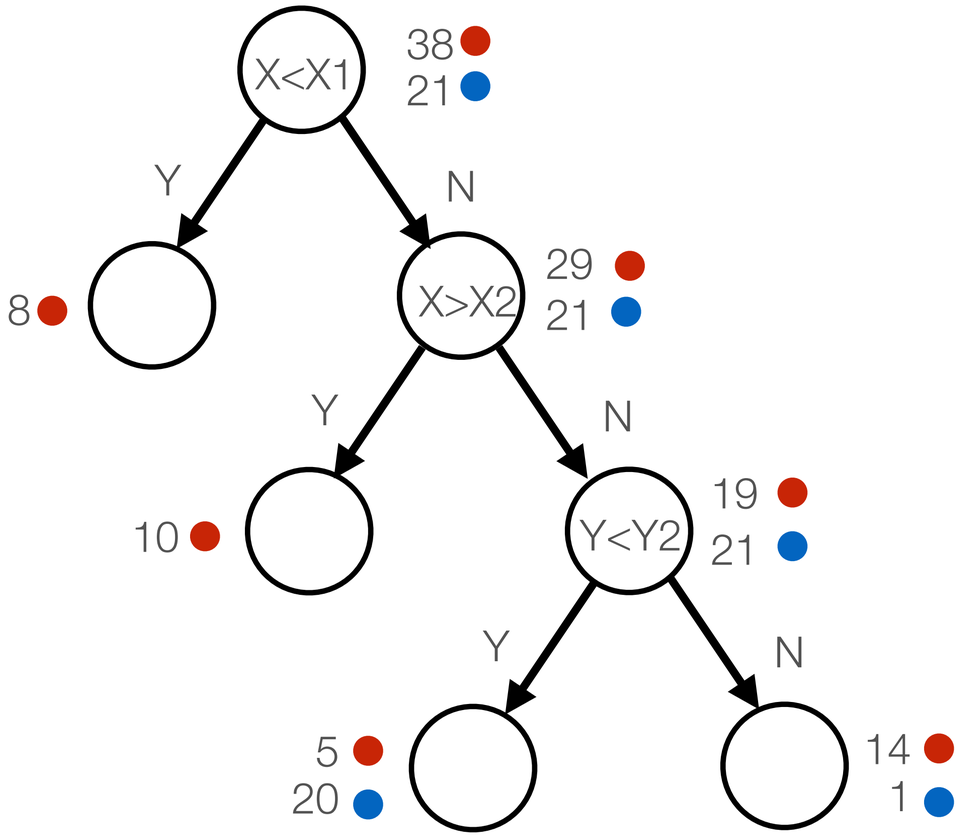
Decision Trees



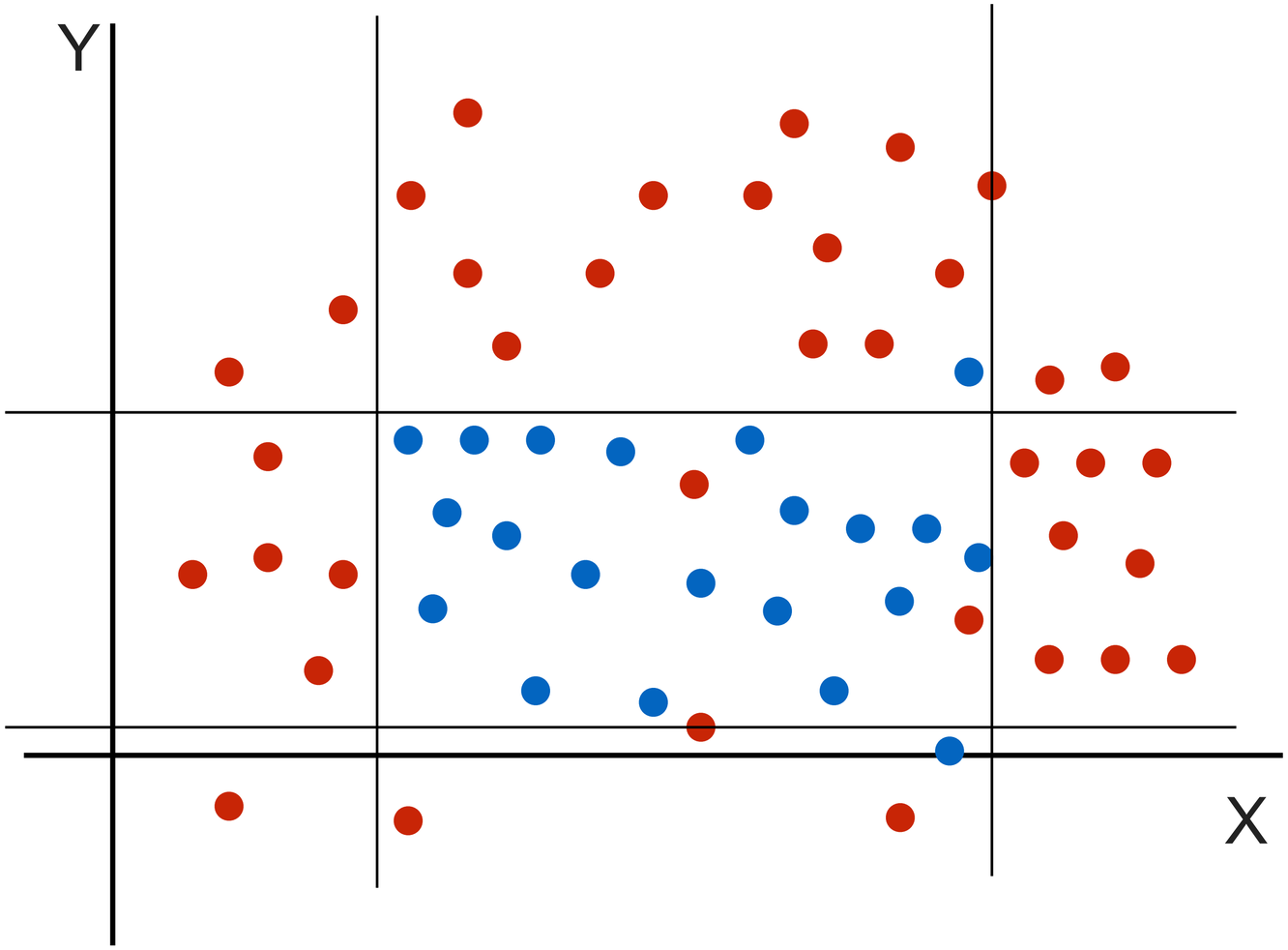


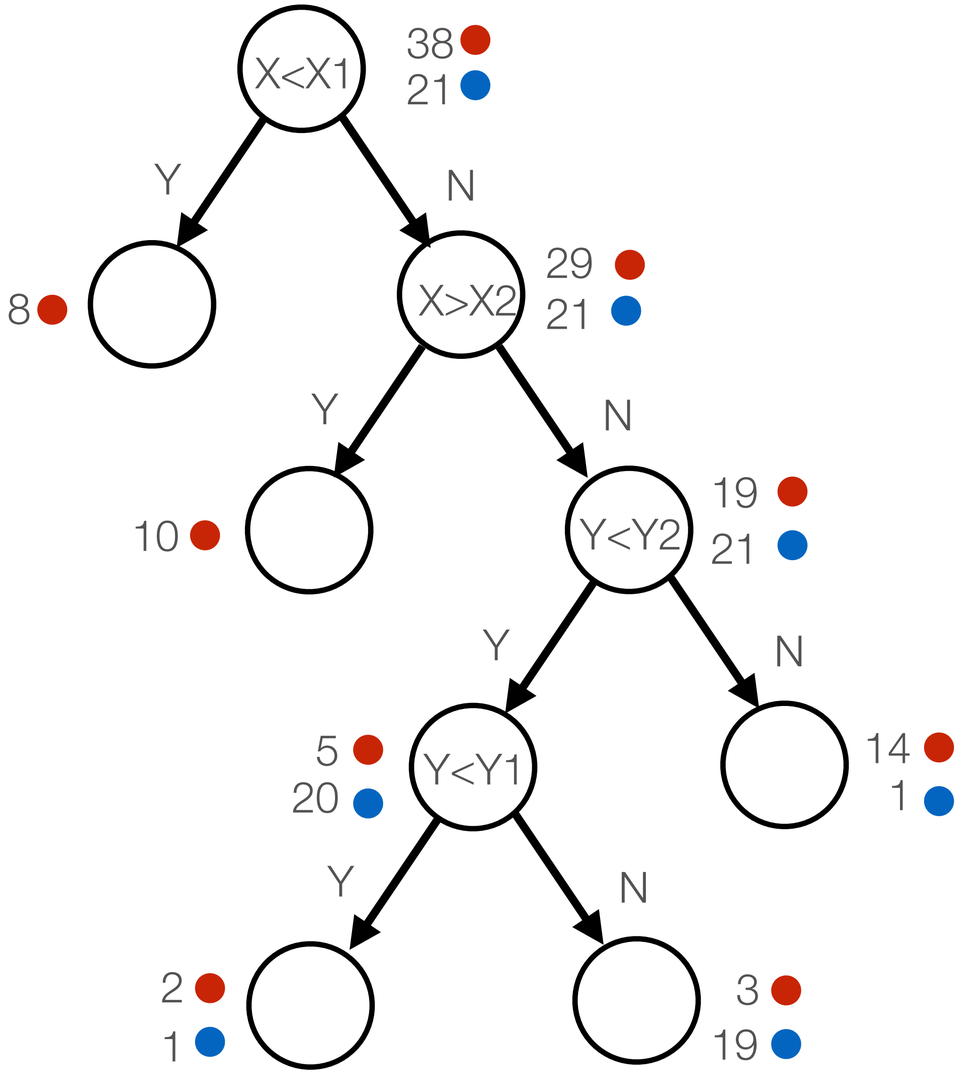
Decision Trees





Decision Trees





@astronomiaOAN

AstronomiaOAN

@astronomiaOAN