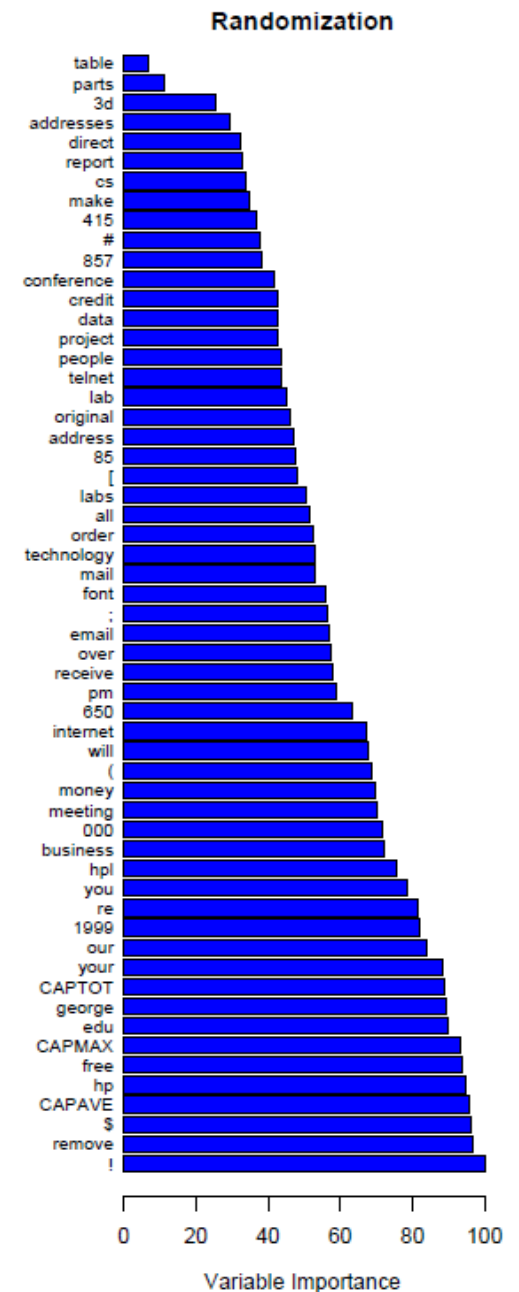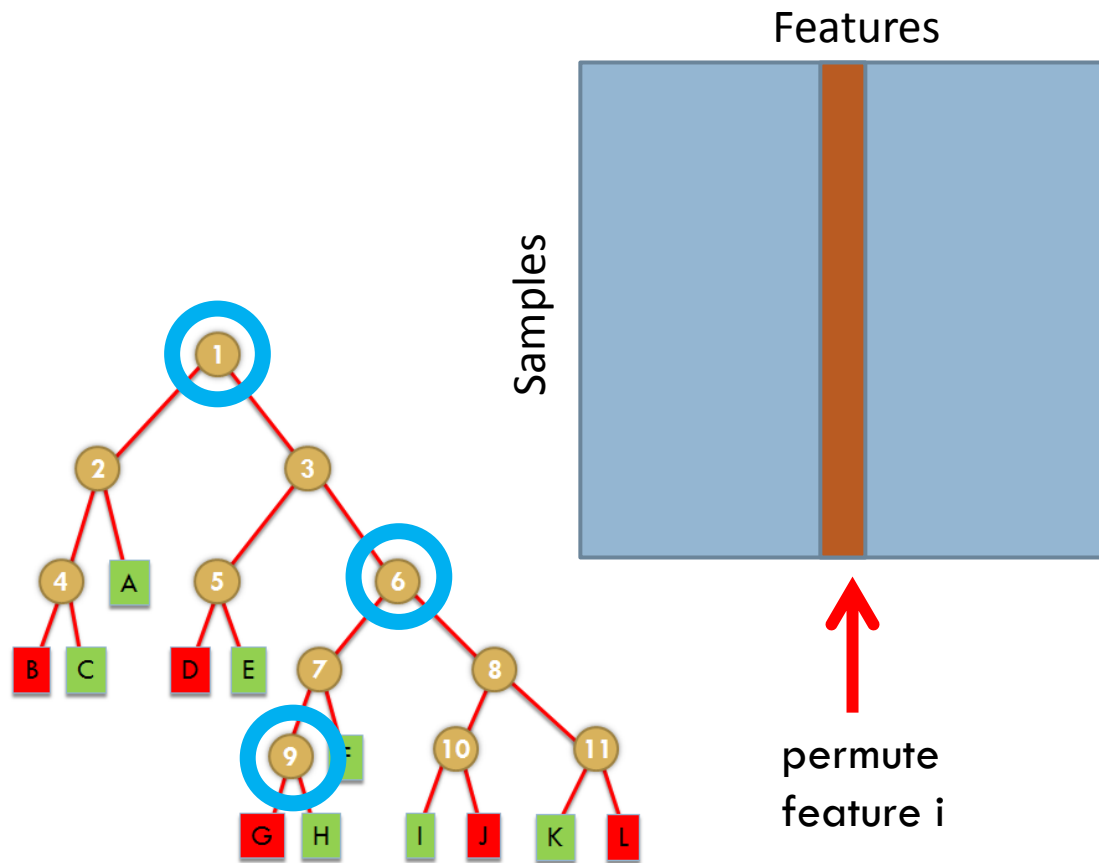Thomas Fuchs (JPL, Caltech)
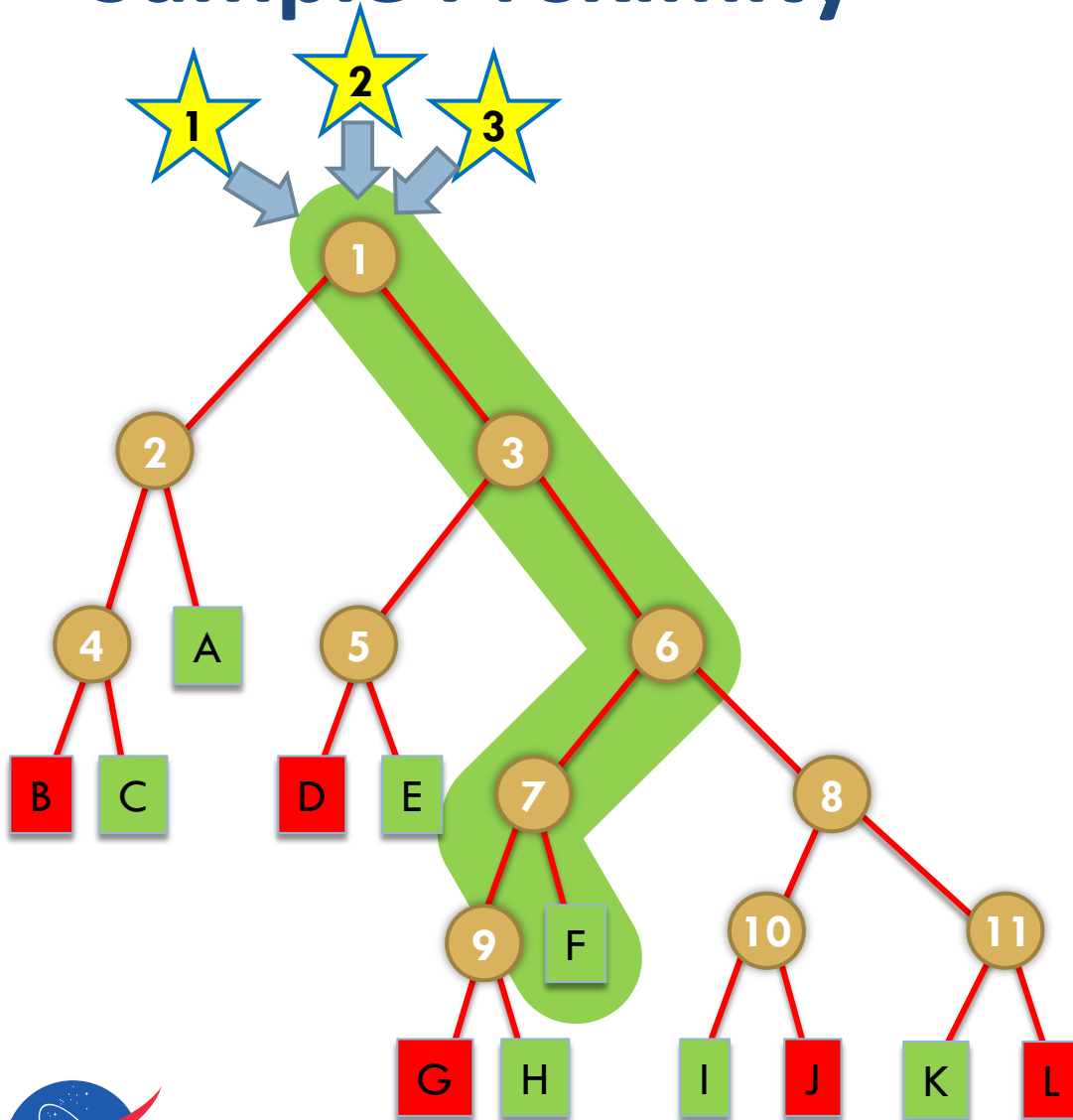# Random Forest Properties

Random Forest Properties

1. Variable Importance

2. Sample Proximity

3. Mixture of Variable Types

4. Parallelization
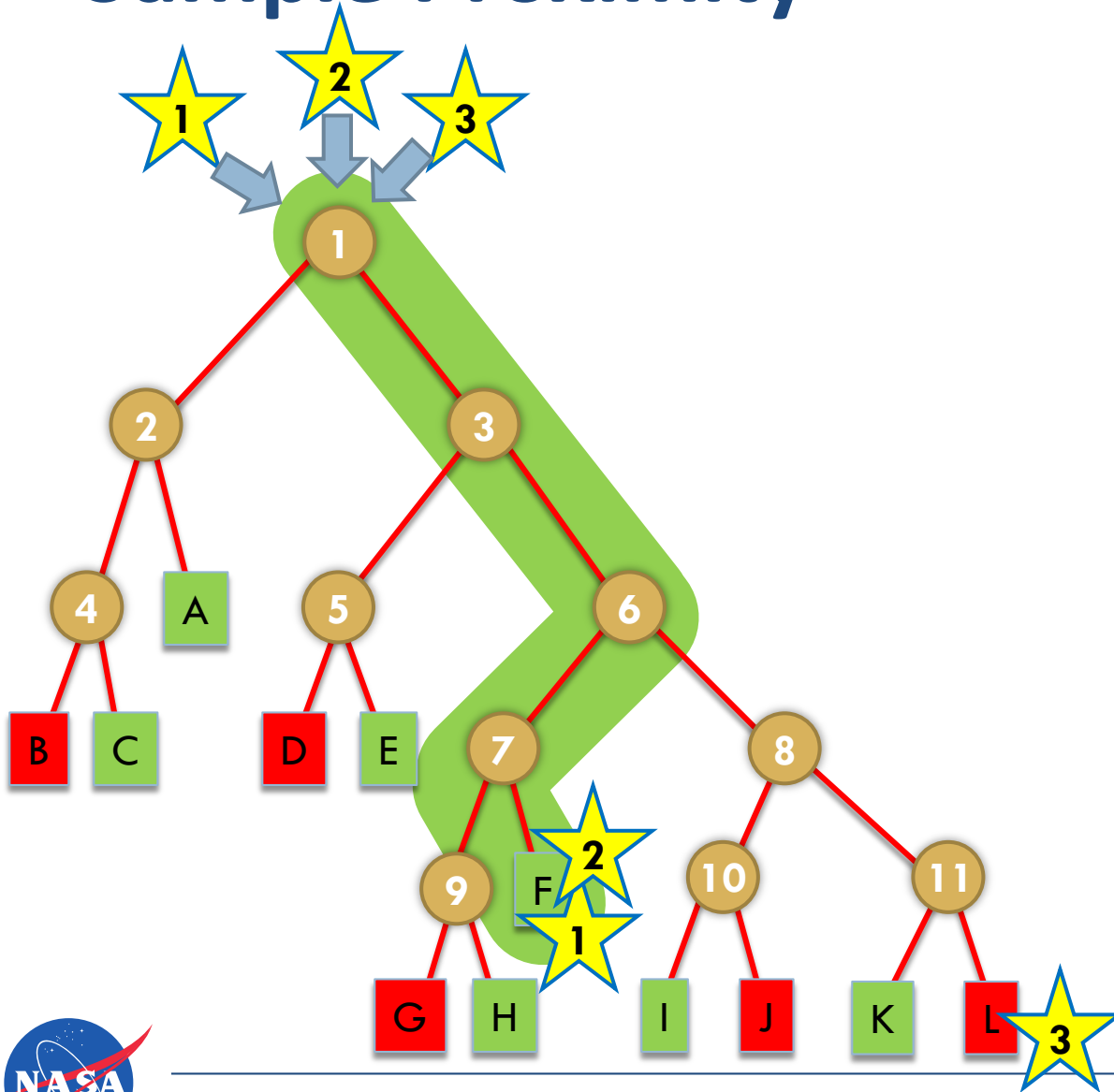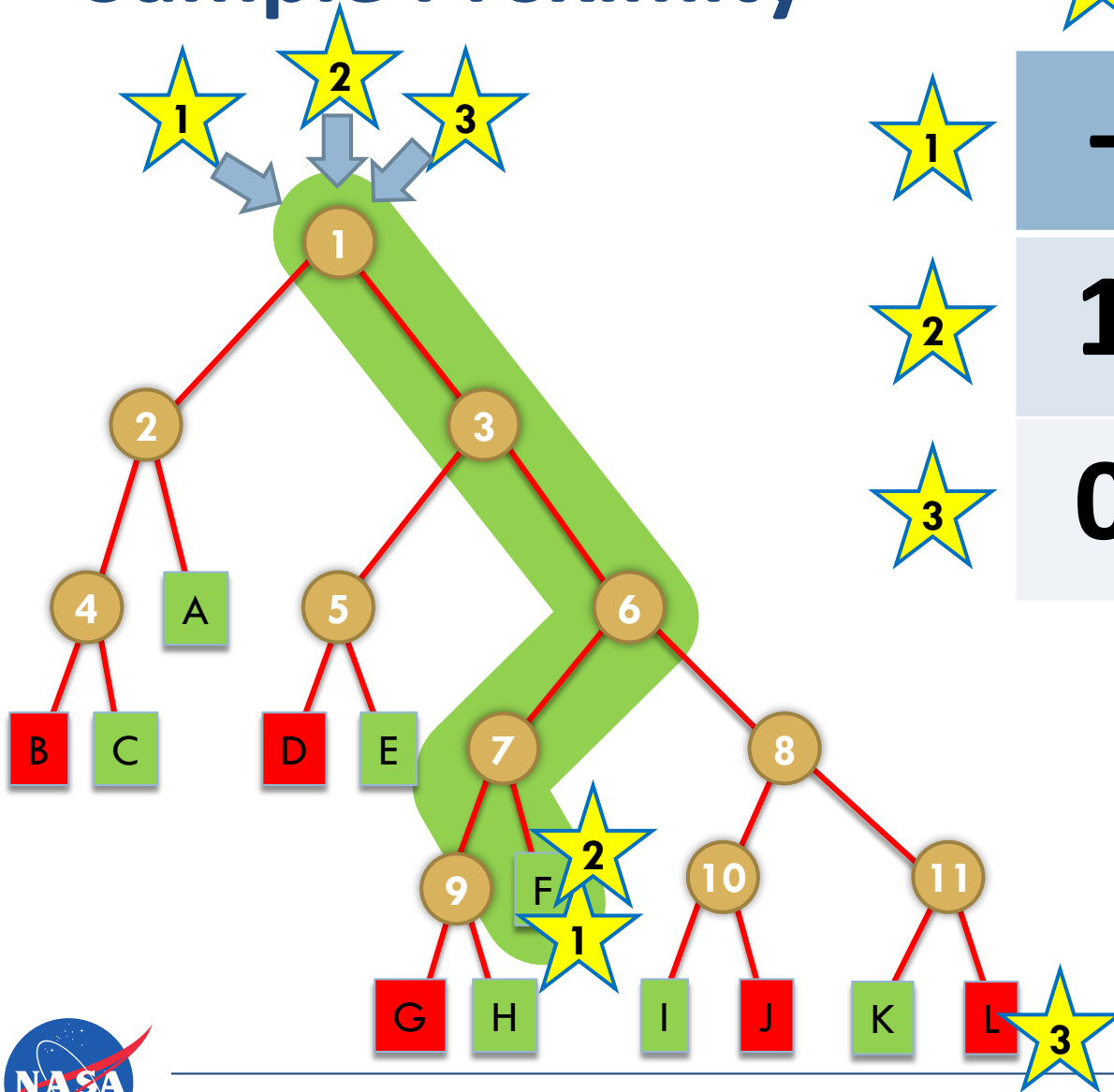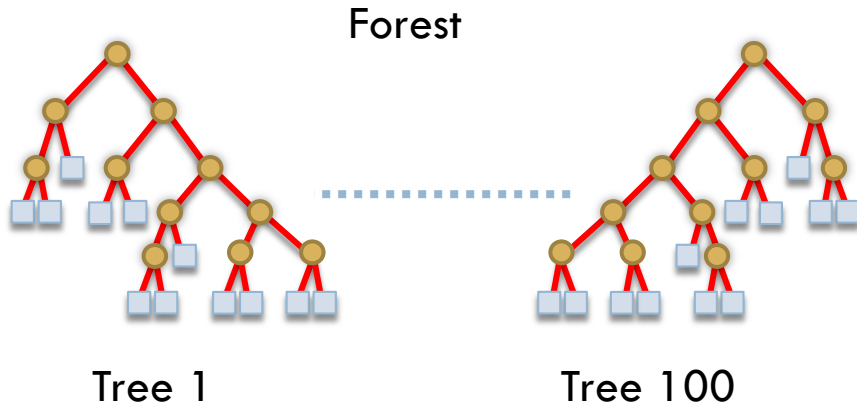
Variable Importance

Features

Samples

permute
feature i

Randomization

Variable Importance

Sample Proximity

Sample Proximity

Sample Proximity

Proximity Matrix

	⭐1	⭐2	⭐3
⭐1	-	1	0
⭐2	1	-	0
⭐3	0	0	-

Sample Proximity

Forest

Tree 1

Tree 100

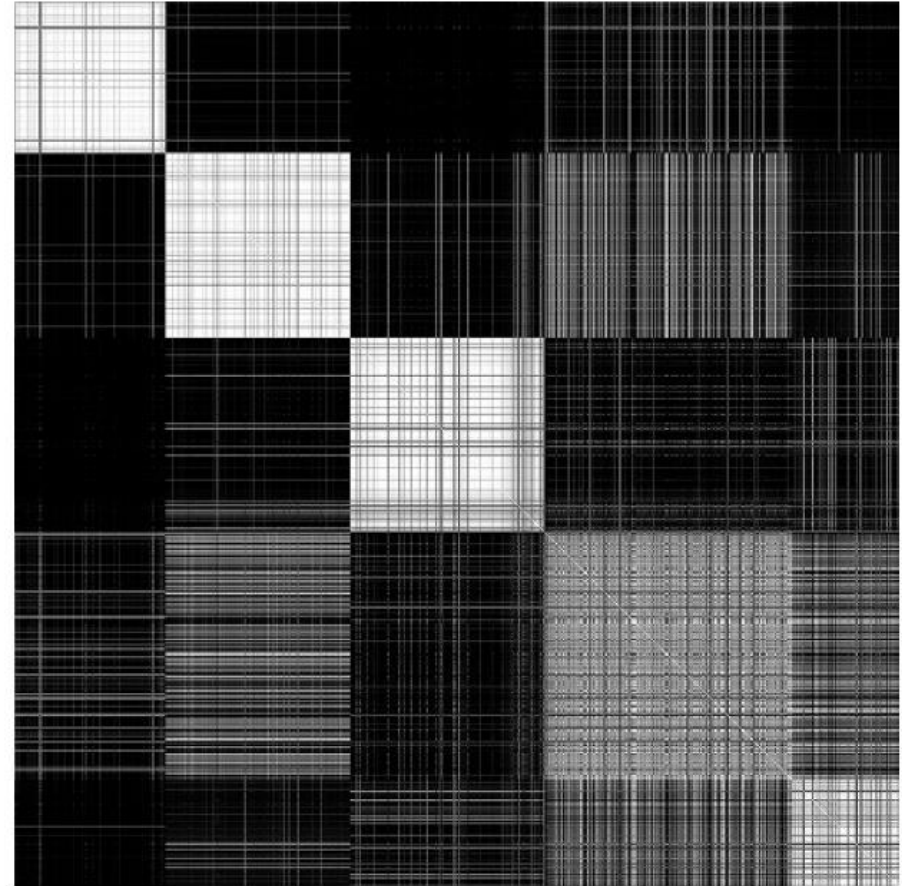|	⭐1	⭐2	⭐3
⭐1	-	83	3
⭐2	83	-	11
⭐3	3	11	-

Proximity Matrix
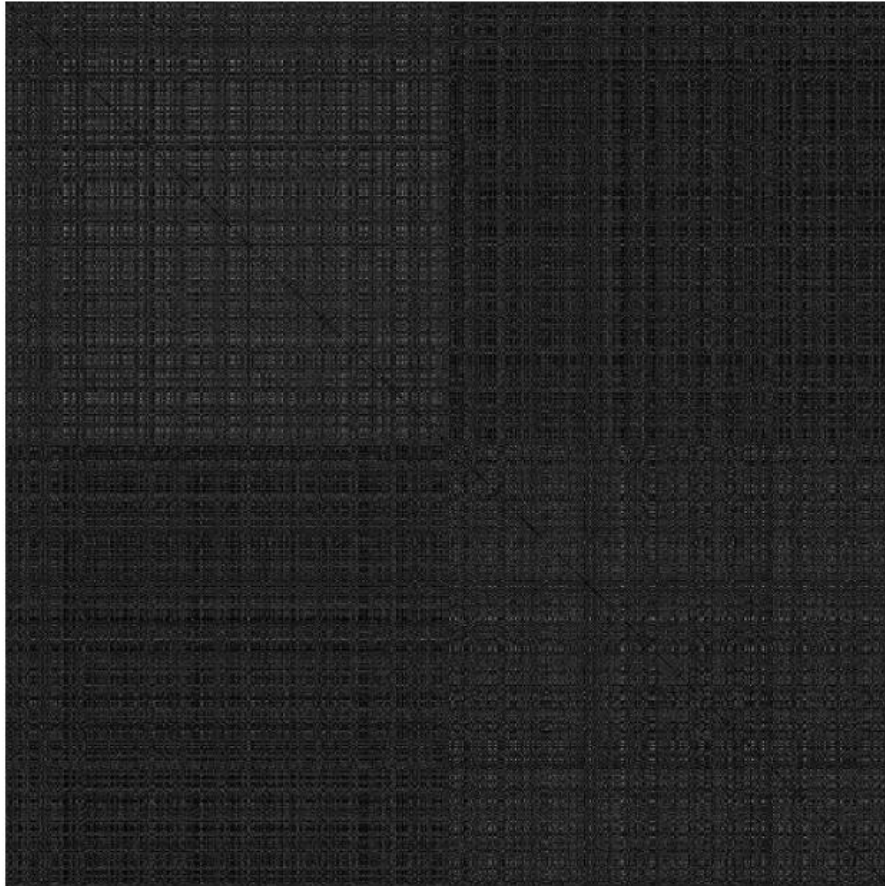
Sample Proximity
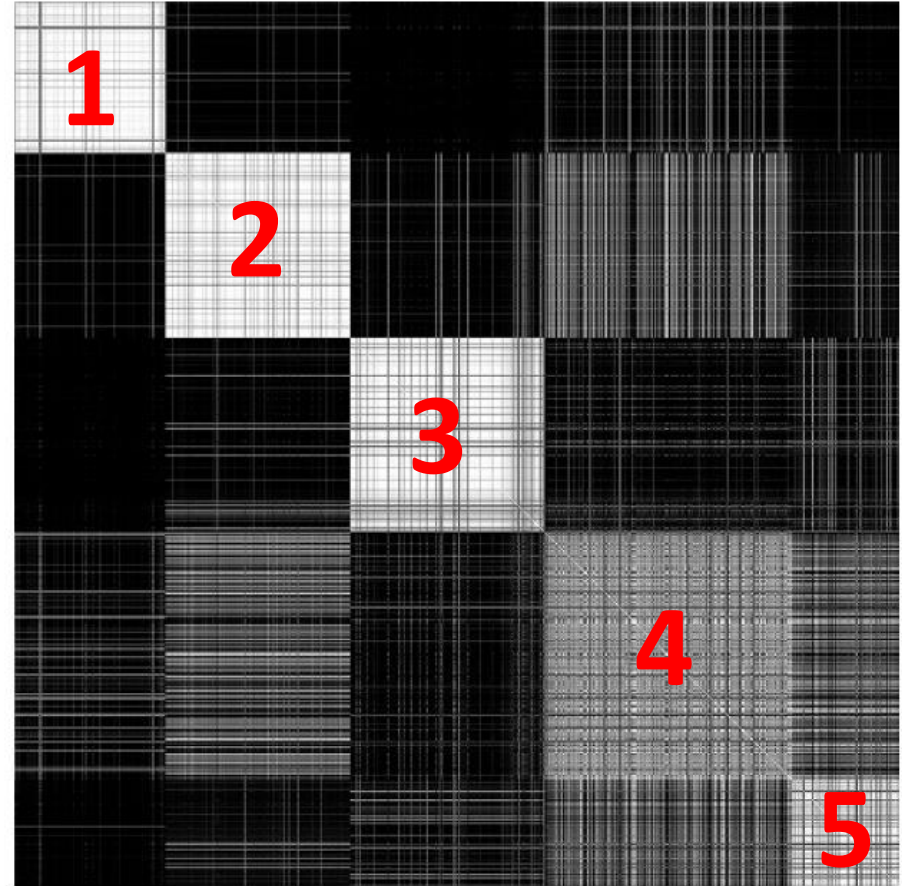

Raw Proximities == Similarity Matrix
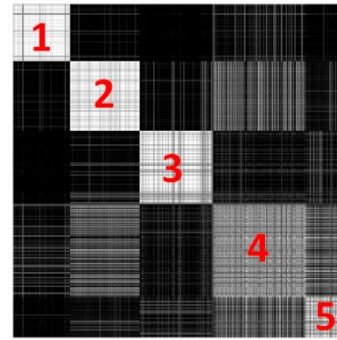
Wishart-Dirichlet Clustering

Sample Proximity

Raw Proximities == Similarity Matrix

Wishart-Dirichlet Clustering
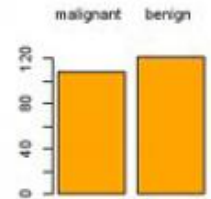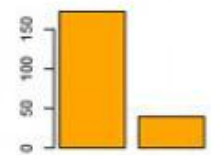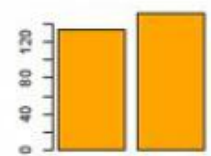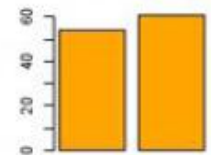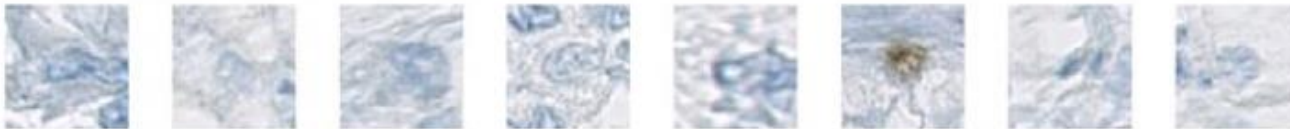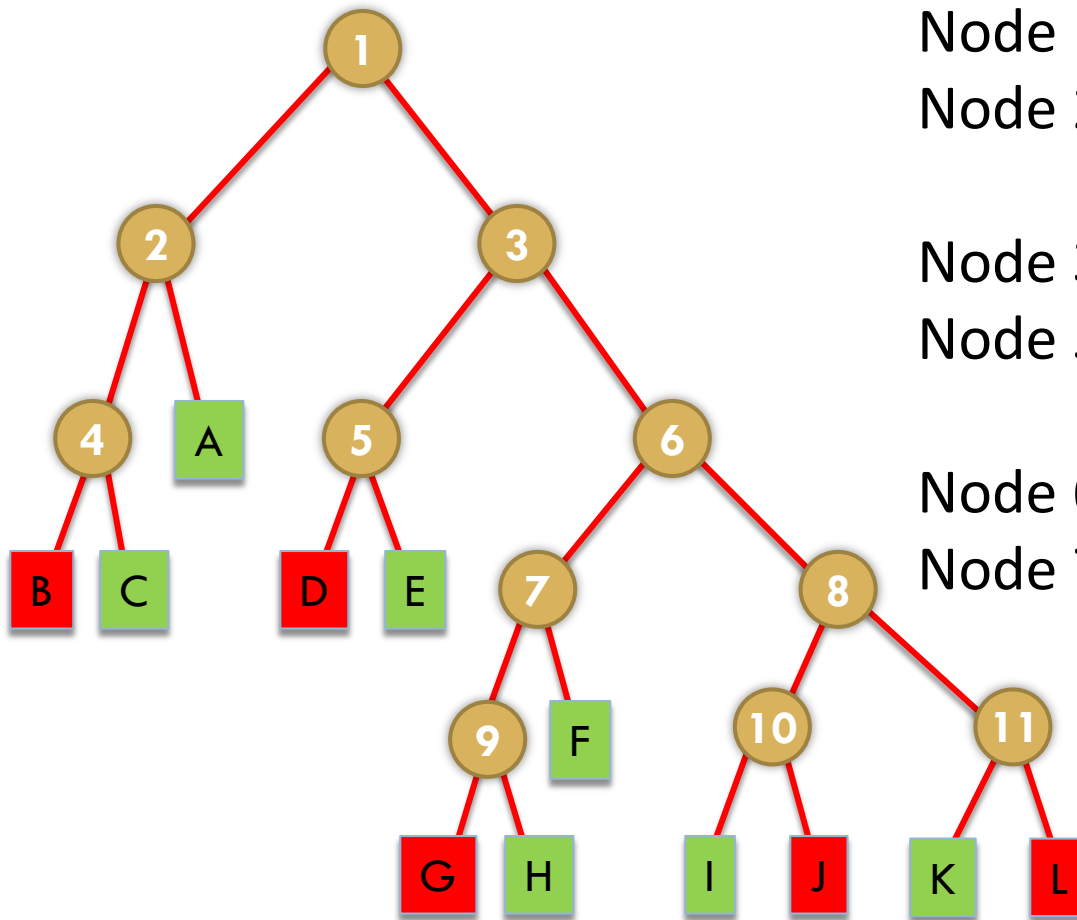
Sample Proximity
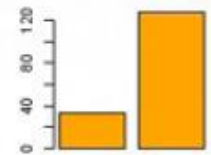
Mixture of Variable Types

Node 1: feature 23 > 0.3
Node 2: feature 11 > 0.1

Node 3: feature 13 == "USA"
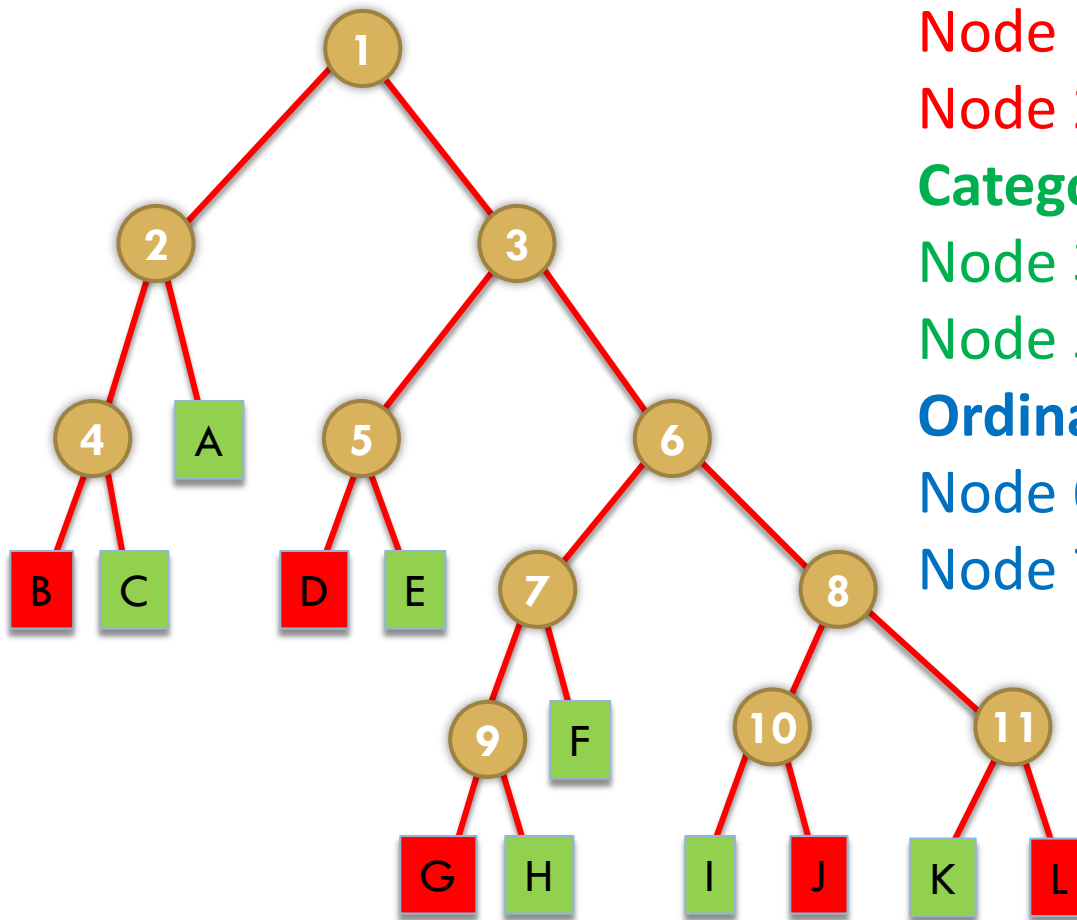Node 5: feature 66 in {LA, NYC, GRZ}

Node 6: feature 27 in < 512
Node 7: feature 27 in > 3

Mixture of Variable Types

Continuous:

Node 1: feature 23 > 0.3

Node 2: feature 11 > 0.1

Categorical:

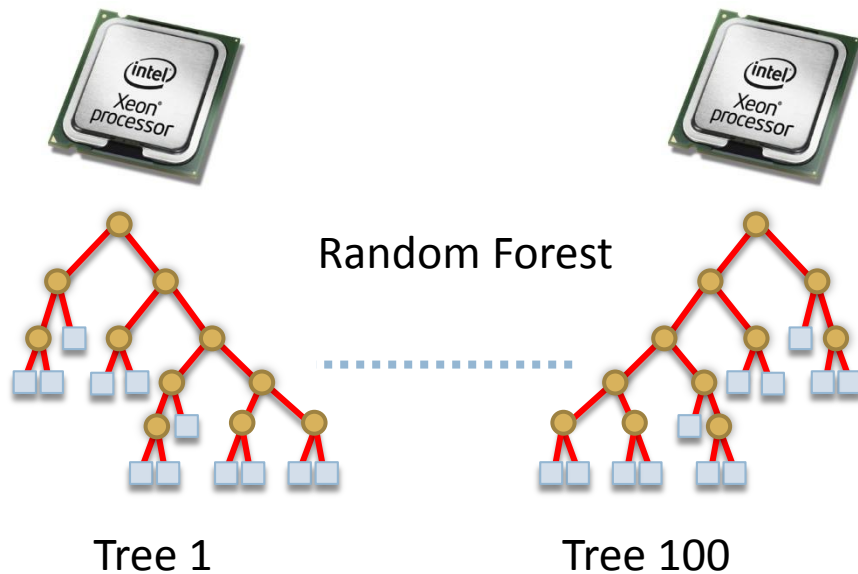Node 3: feature 13 == "USA"

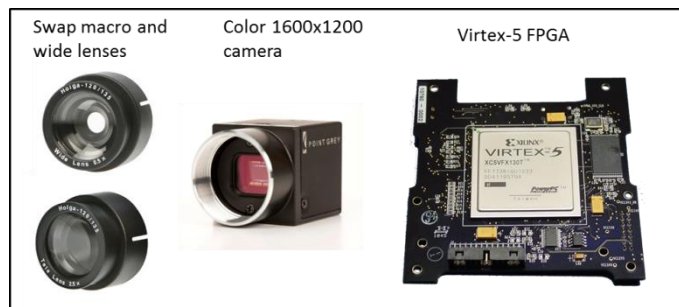Node 5: feature 66 in {LA, NYC, GRZ}

Ordinal:

Node 6: feature 27 in < 512

Node 7: feature 27 in > 3

Parallel Training and Testing

Random Forest

Tree 1 .... Tree 100

In addition optimized implementations for GPU and FPGA exist.

Example:
JPL's Texturecam

Swap macro and wide lenses
Color 1600x1200 camera
Virtex-5 FPGA

What we'll build

FPGA Computer augmenting a camera

Entire apparatus fits in a "shoebox"

A

B