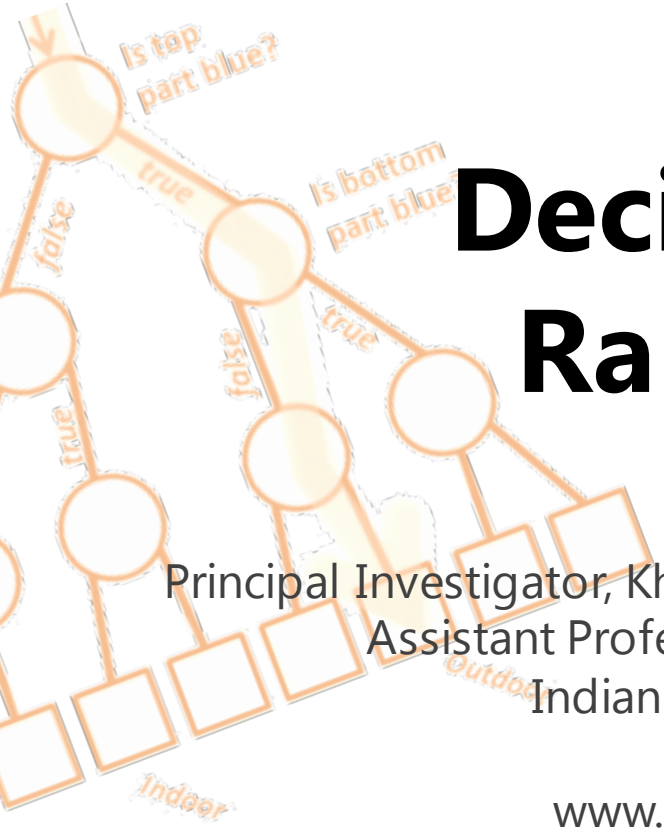


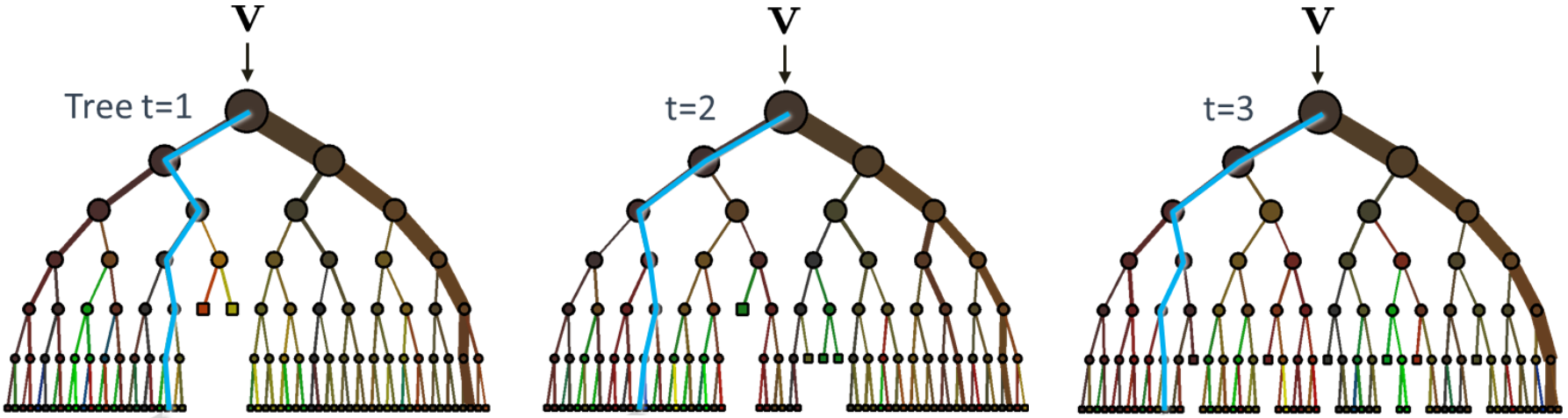
Decision Trees and Random Forests

Dr. Debdoot Sheet
Principal Investigator, Kharagpur Learning, Imaging and Visualization Group
Assistant Professor, Department of Electrical Engineering
Indian Institute of Technology Kharagpur

www.facweb.iitkgp.ernet.in/~debdoot/

NOT ABOUT WALKING IN A FOREST

IS ALL ABOUT



Overview

- Historical Perspective
- Decision Tree
- Random Forest
- Computational Complexity
- Variable Importance

Historical Perspective

Decision Trees

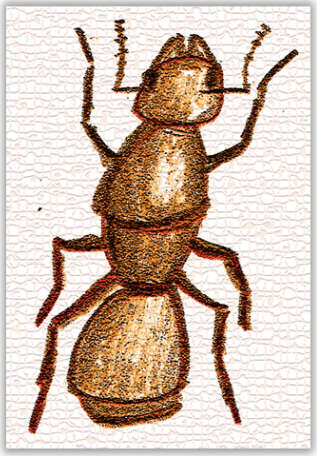
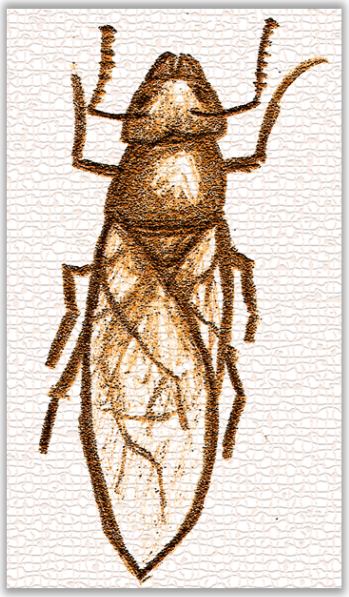
- L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Chapman and Hall/CRC (SIAM), **1984**.
- J. R. Quinlan, *C4.5: Programs for Machine Learning*. **1993**.

Random Forests

- Y. Amit and D. Geman., "Shape quantization and recognition with randomized trees," *Neural Computation,* vol. 9, pp. 1545–1588, **1997**.
- T. K. Ho, "The random subspace method for constructing decision forests," *IEEE T-PAMI,* vol. 20, no. 8, pp. 832–844, **1998**.
- L. Breiman, "Random forests," *Machine Learning,* vol. 45, no. 1, pp. 5–32, **2001**.

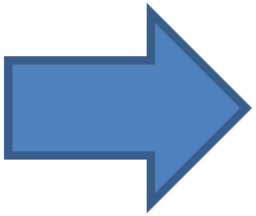
DECISION TREE

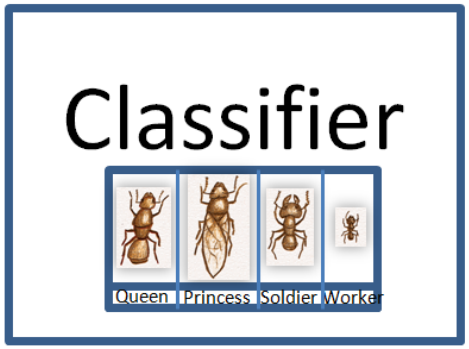
Problem Statement

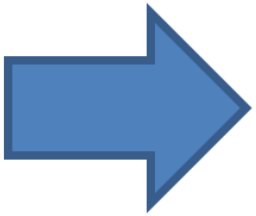
			
Queen	Princess	Soldier	Worker

Formica rufa (Red wood ant)

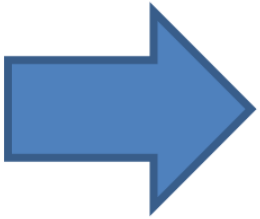
Classification vs. Regression

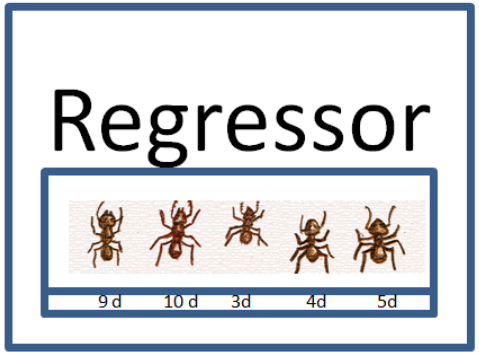


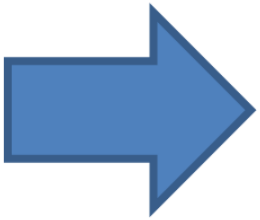




Worker

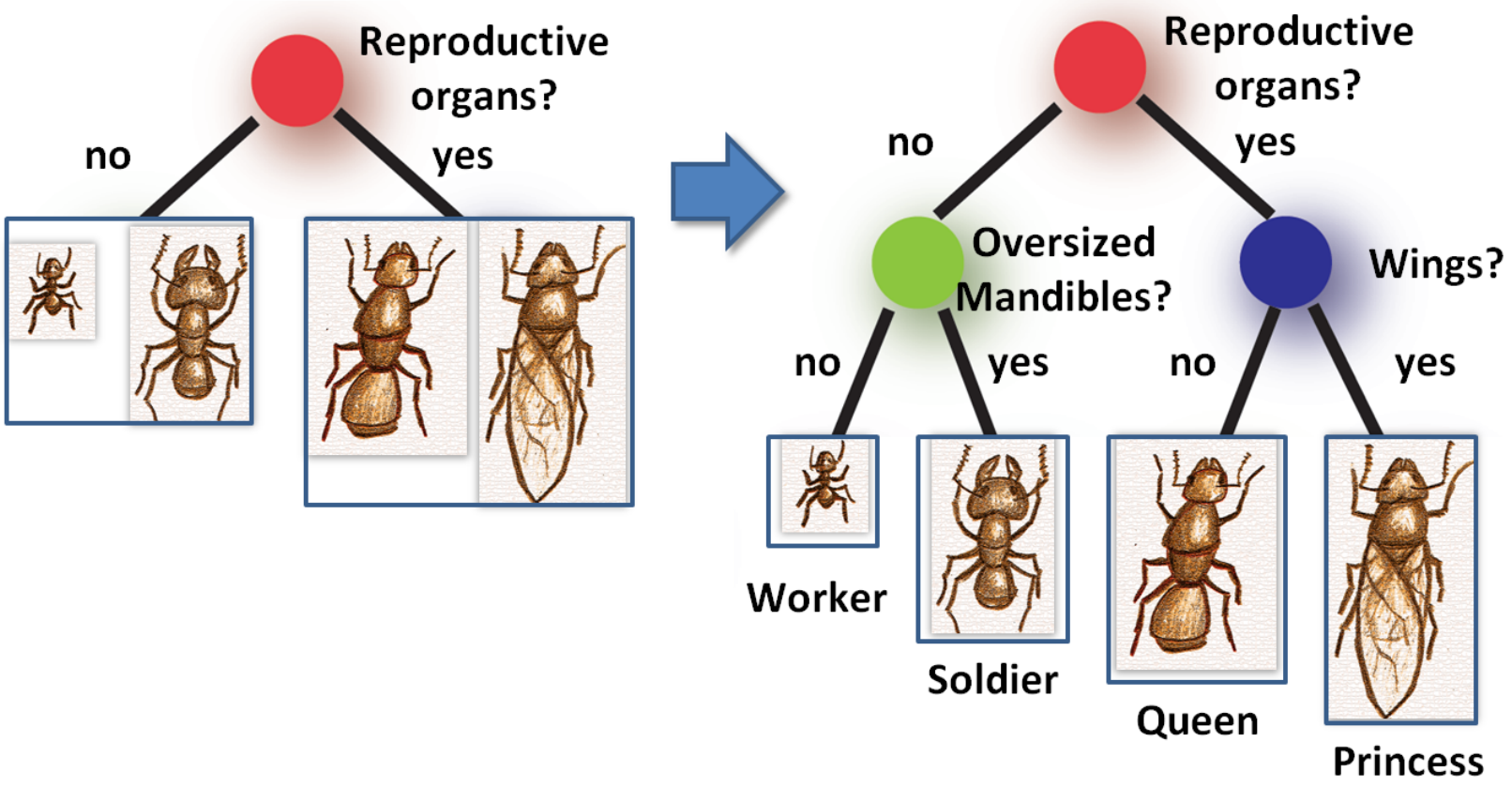




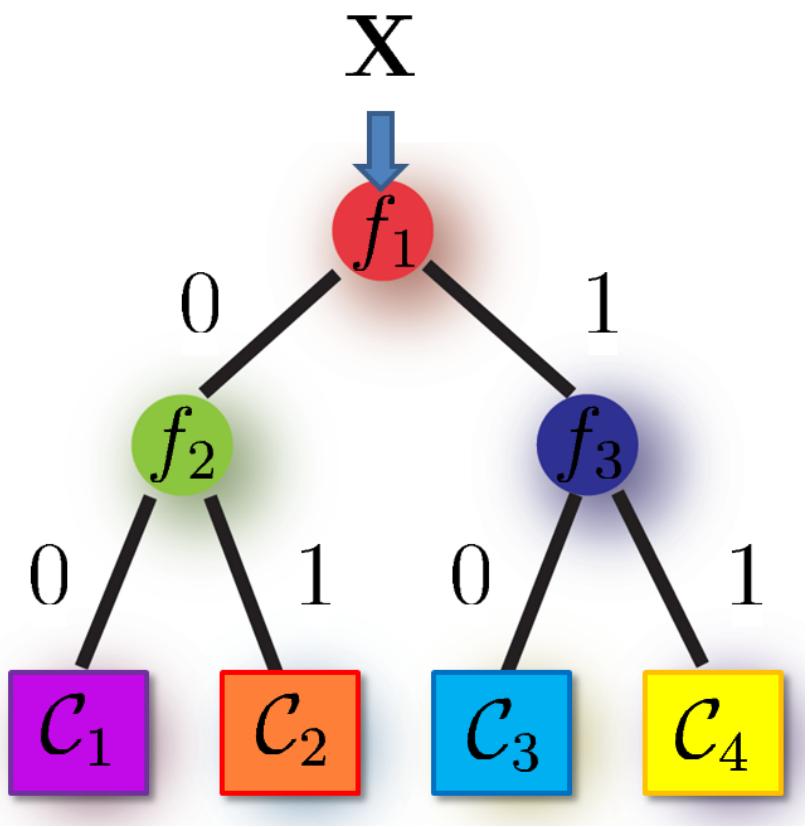


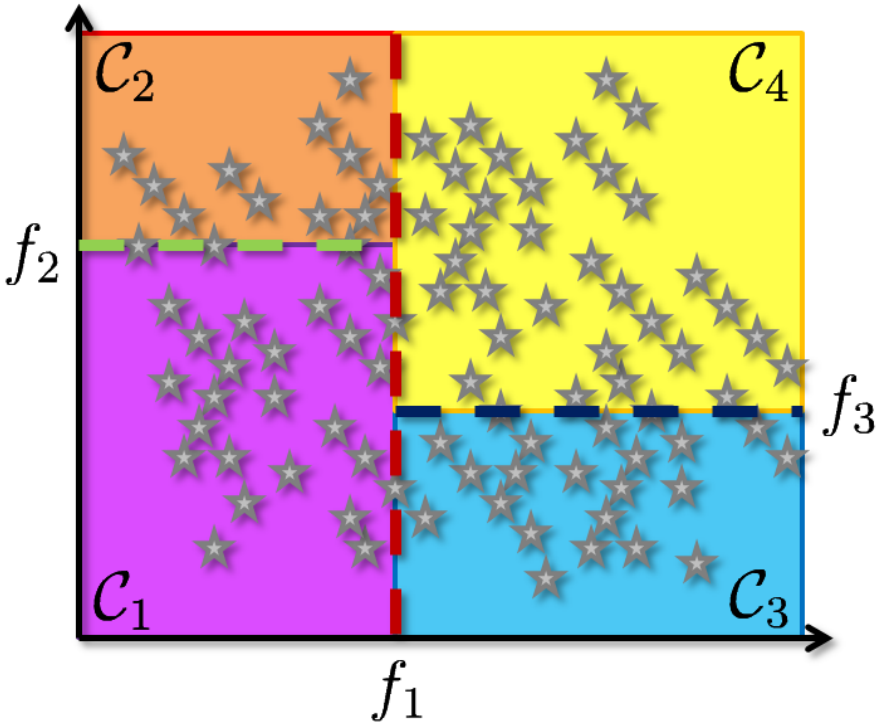
6 days

Decision Tree

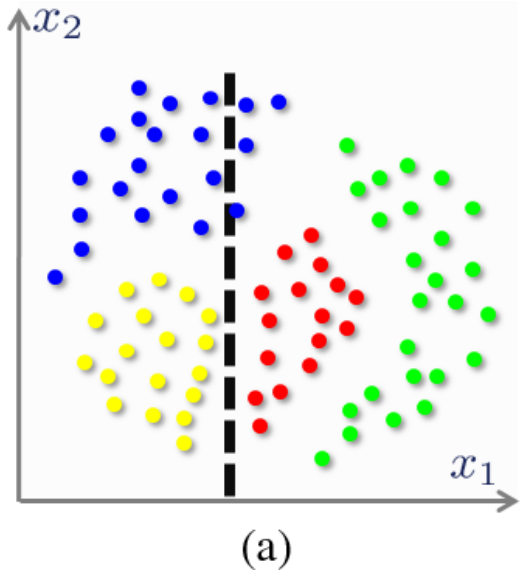


Forming a Decision Tree



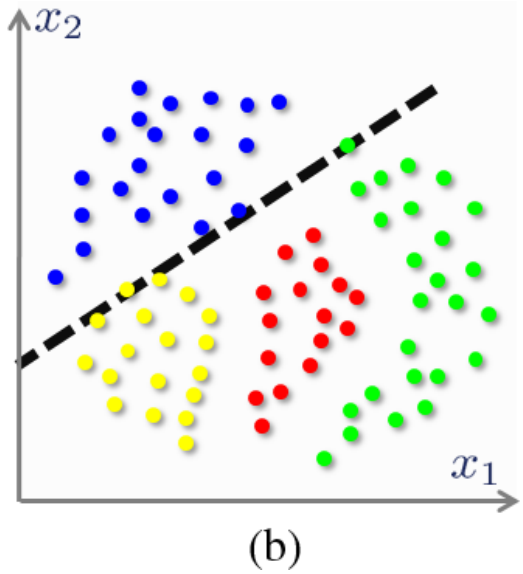


Step 1: Split Function at Node



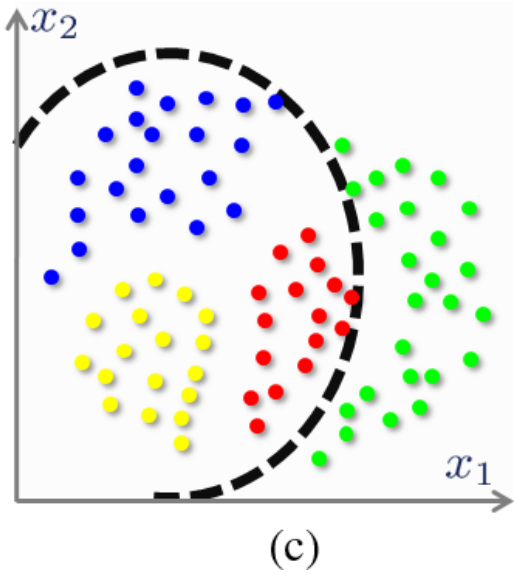
(a)

Axis aligned split



(b)

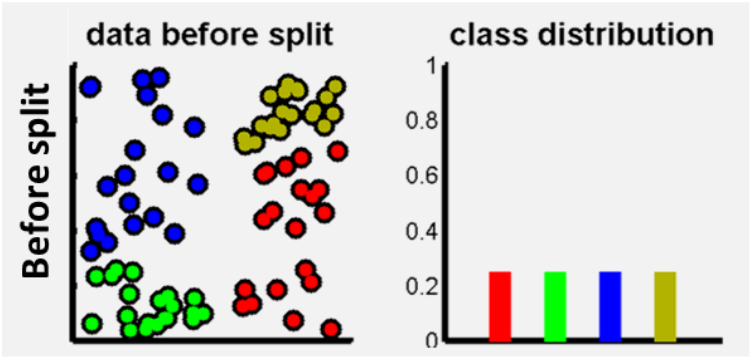
Oblique split



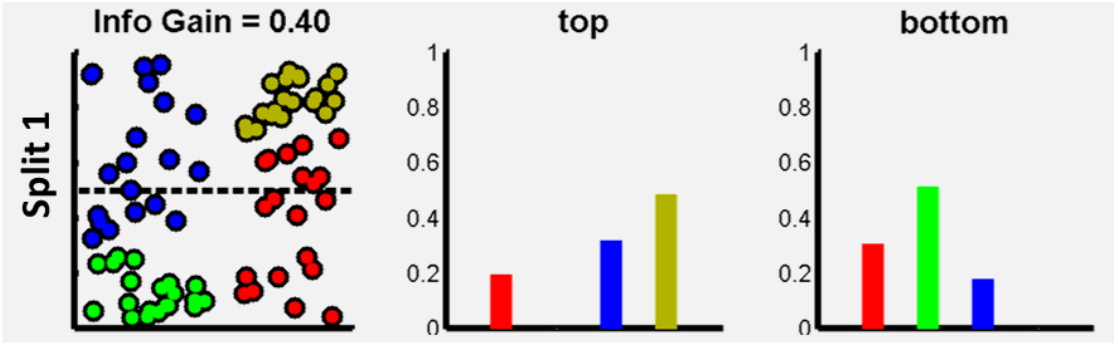
(c)

Polynomial split

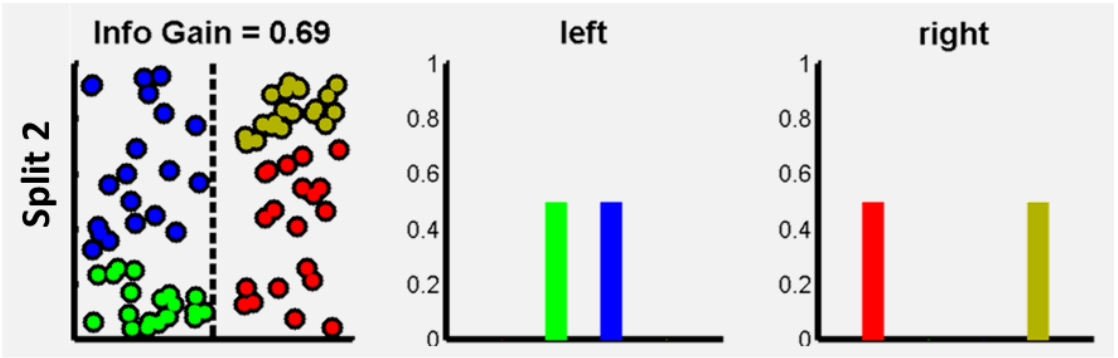
Step 2: Assessing Purity of Split



(a)

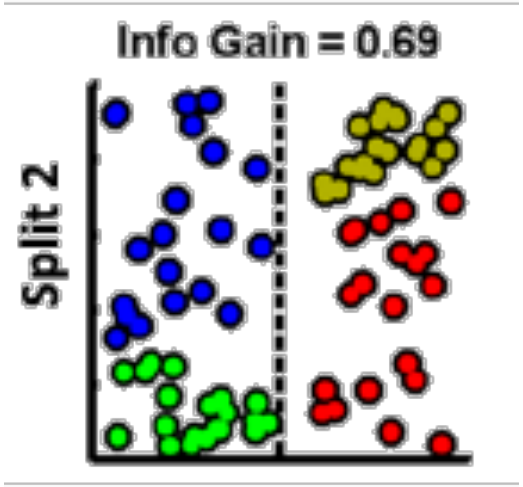


(b)



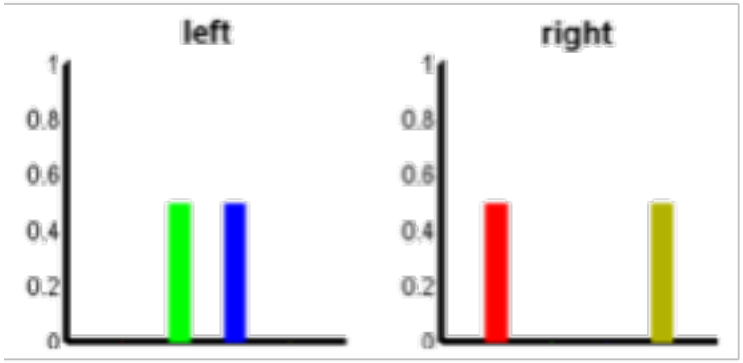
(c)

Cost function for Split Purity



Entropy of class distribution

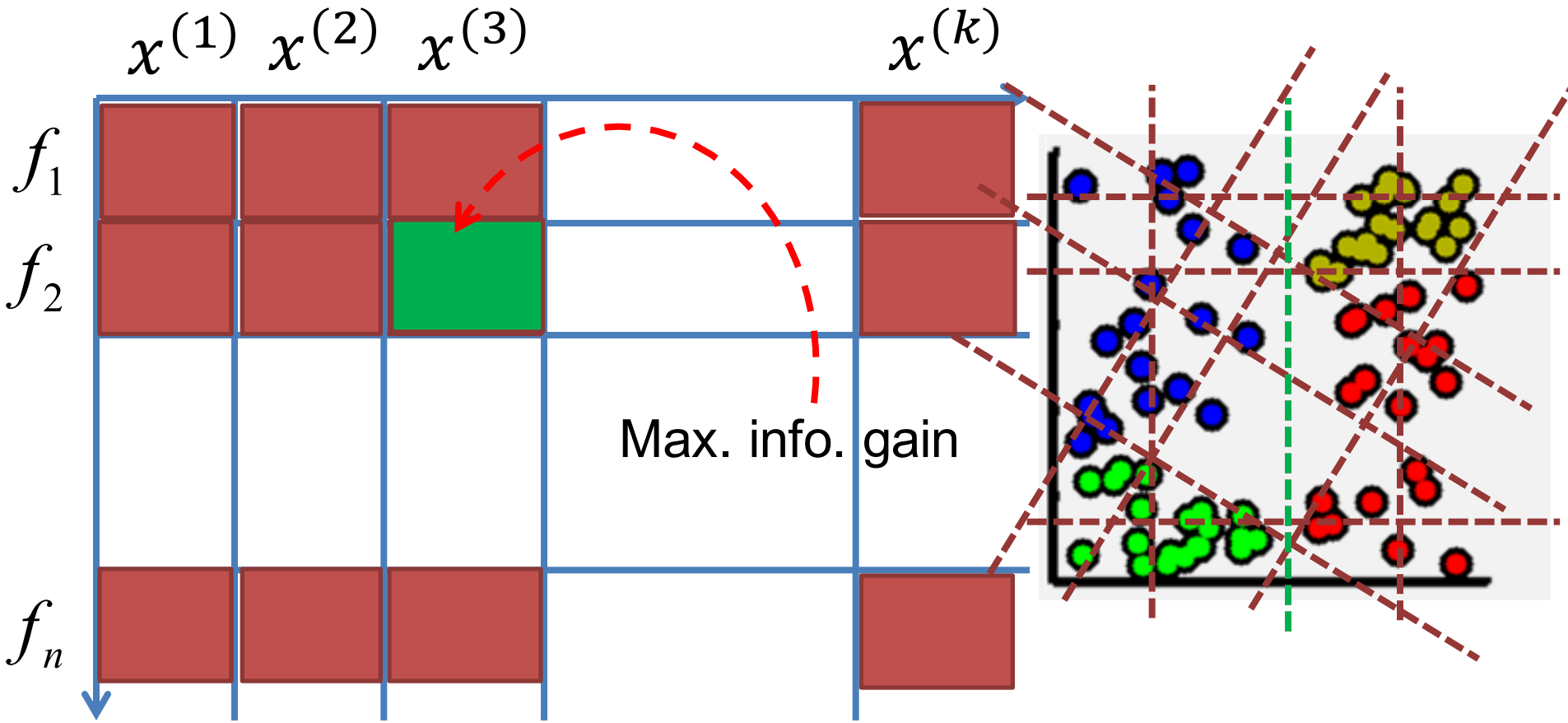
$$H(\mathcal{S}) = -\sum_{c \in \mathcal{C}} p(c) \log(p(c)),$$



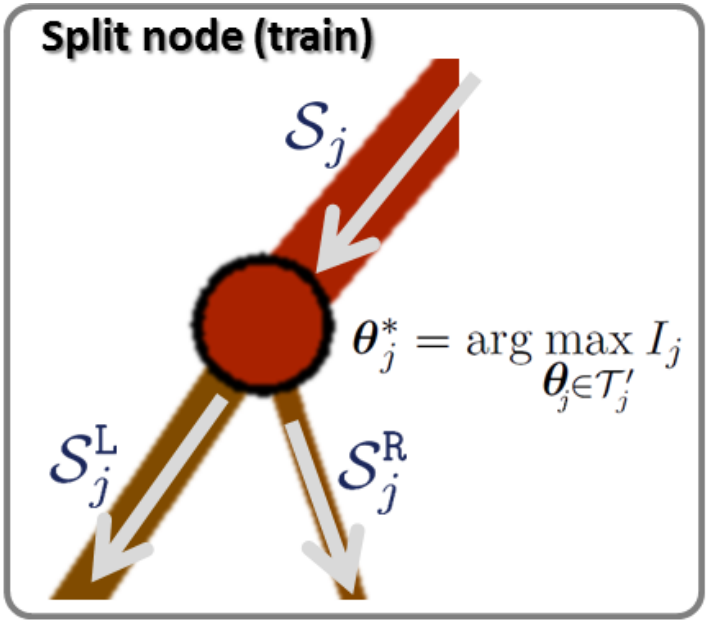
Information Gain

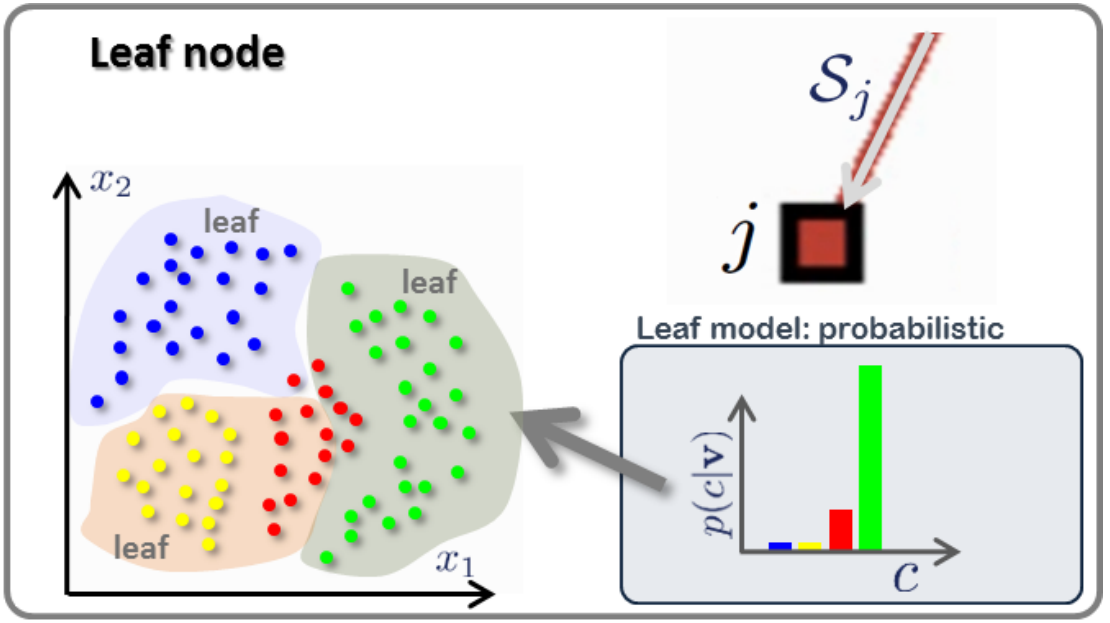
$$I = H(\mathcal{S}) - \sum_{i \in \{\mathrm{L},\mathrm{R}\}} \frac{|\mathcal{S}^i|}{|\mathcal{S}|} H(\mathcal{S}^i).$$

Step 3: Selecting Optimum Split

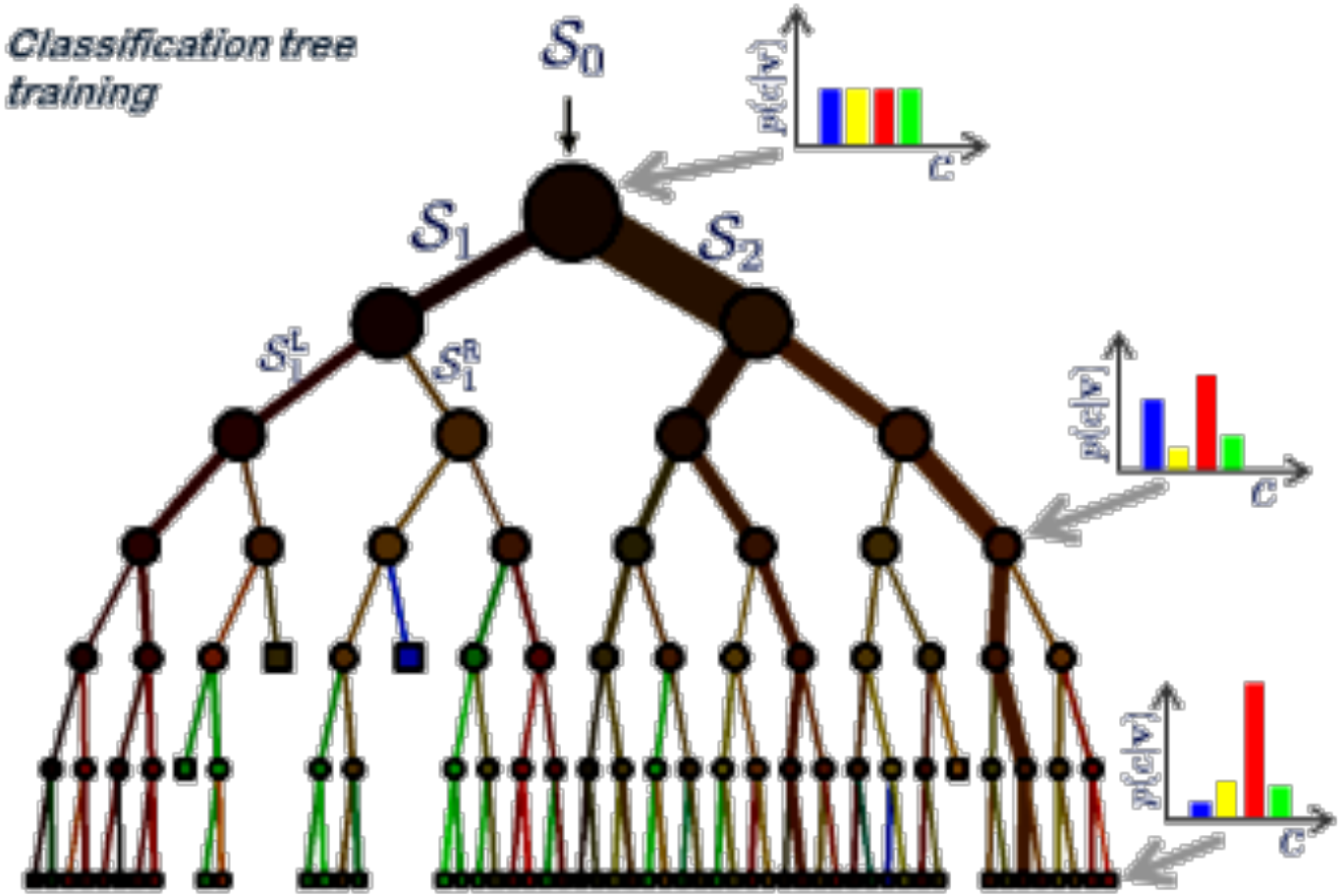


Step 4: Stopping Criteria

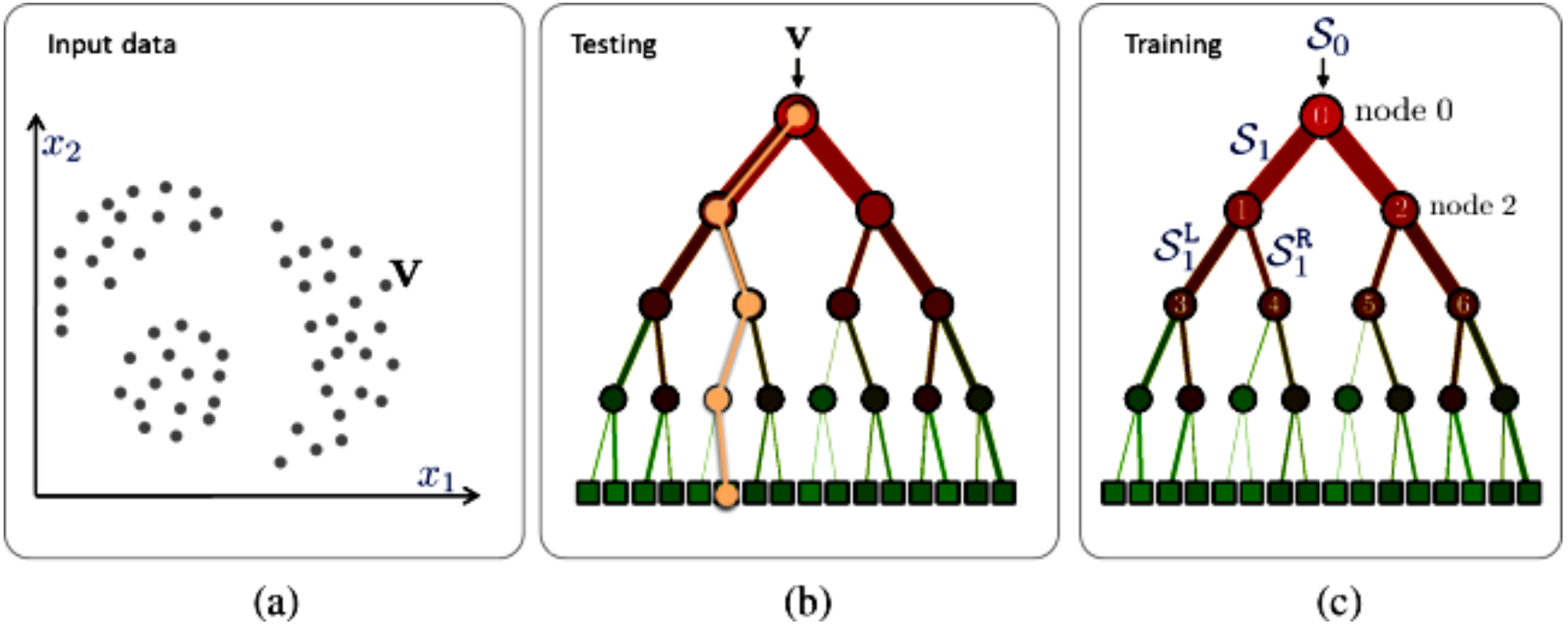




Step 5: Leaf Prediction Model

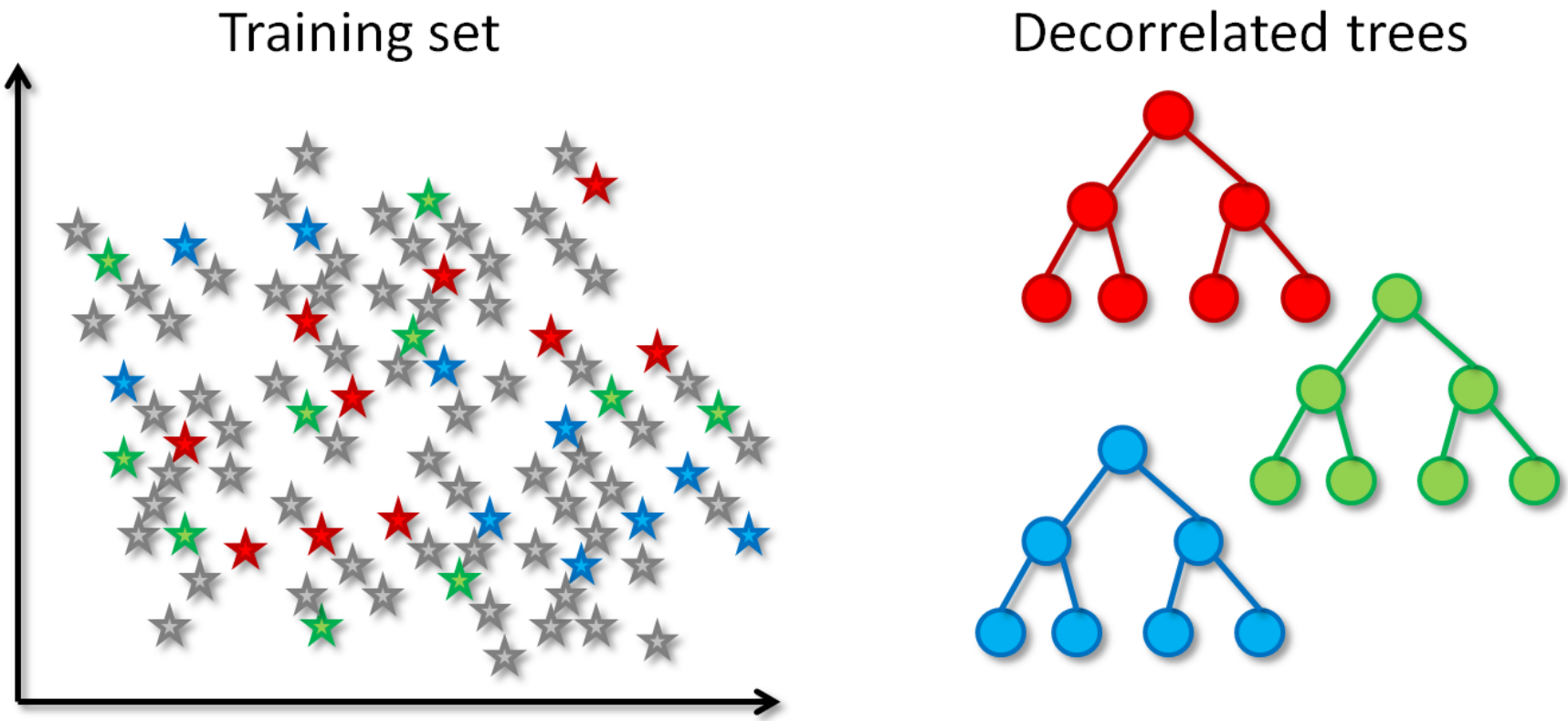


Deploying a Decision Tree



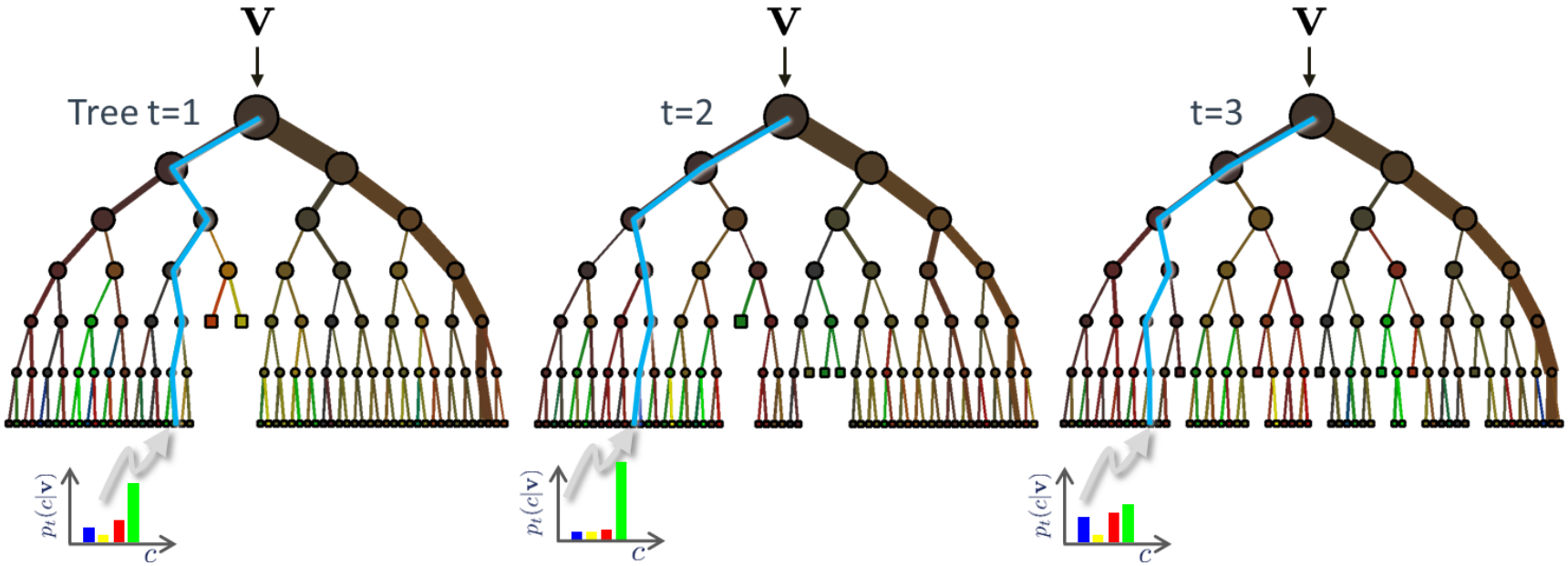
RANDOM FOREST

Growing Multiple Trees in a Forest



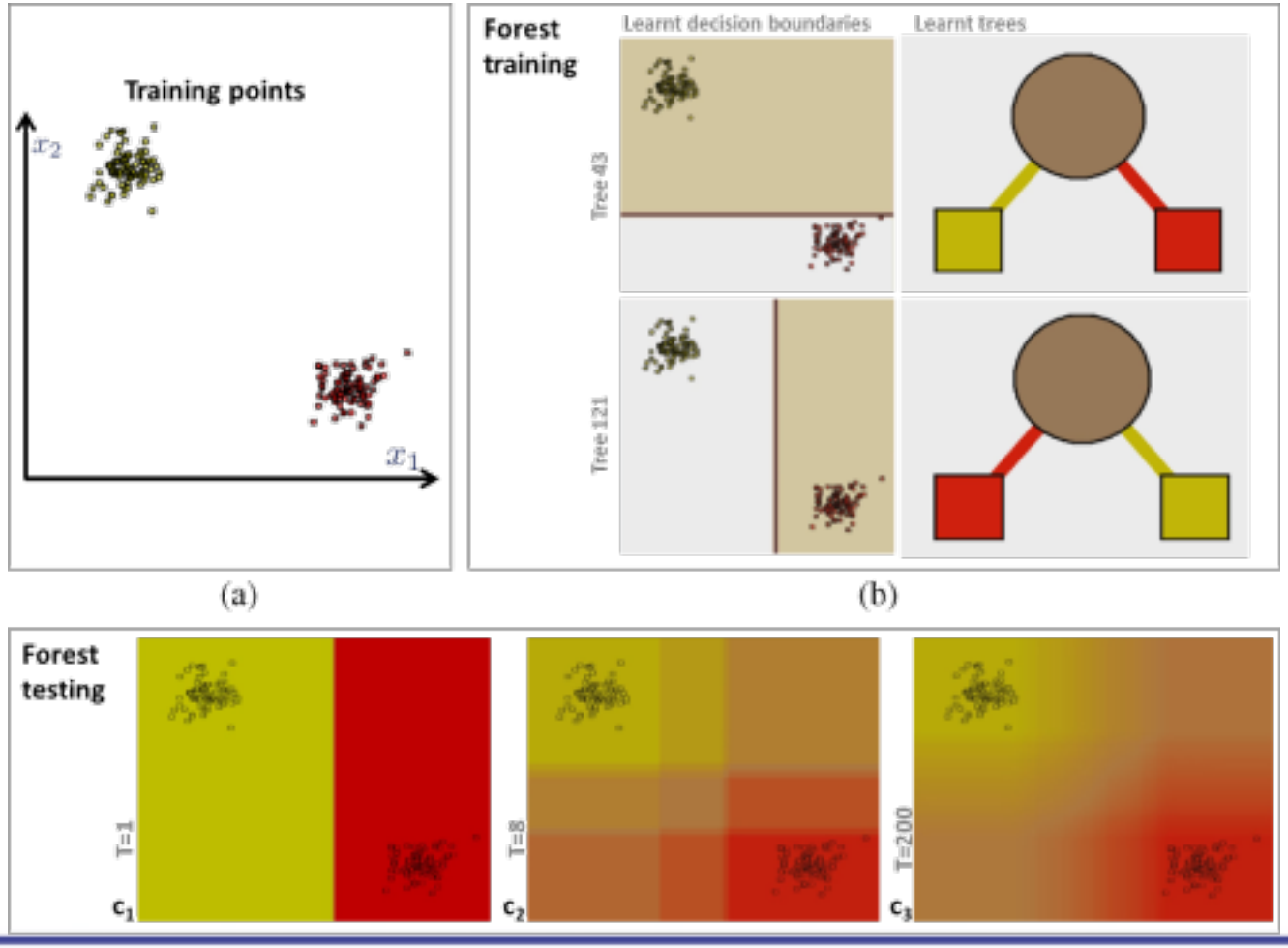
Bagging – Bootstrapped Aggregation

Ensemble Prediction Model

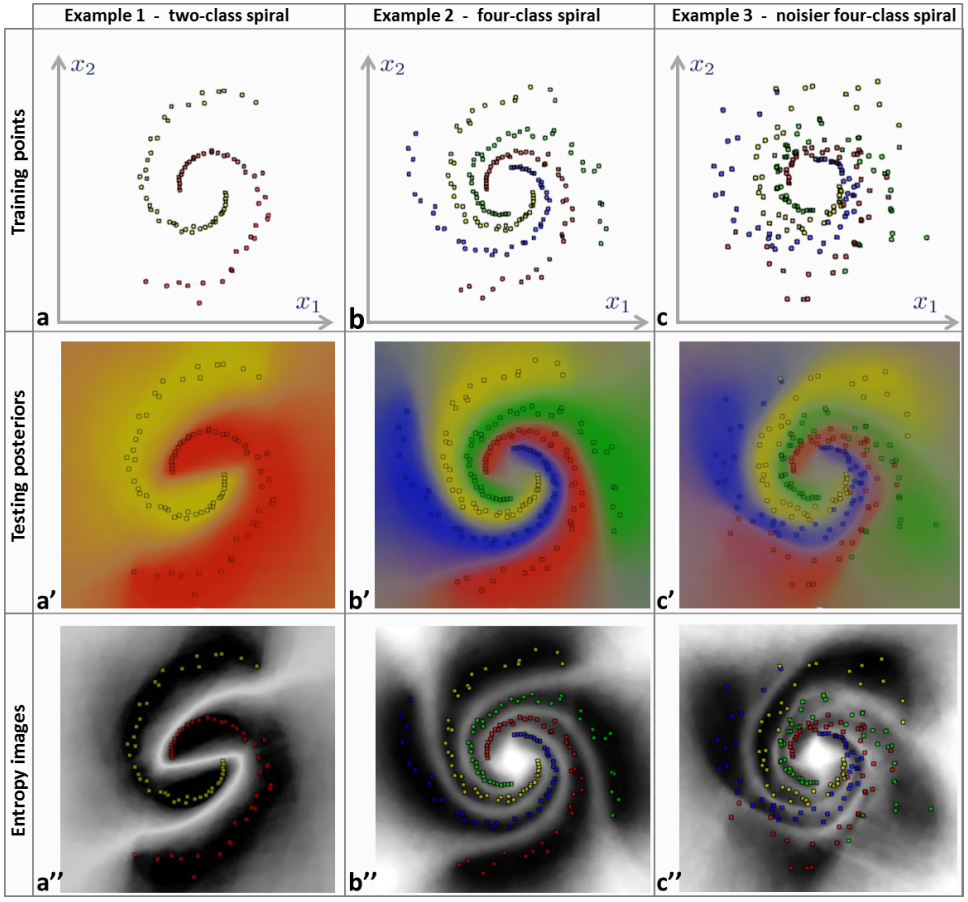


$$p(c|\mathbf{v}) = \frac{1}{T}\sum_{t}^{T} p_t(c|\mathbf{v})$$

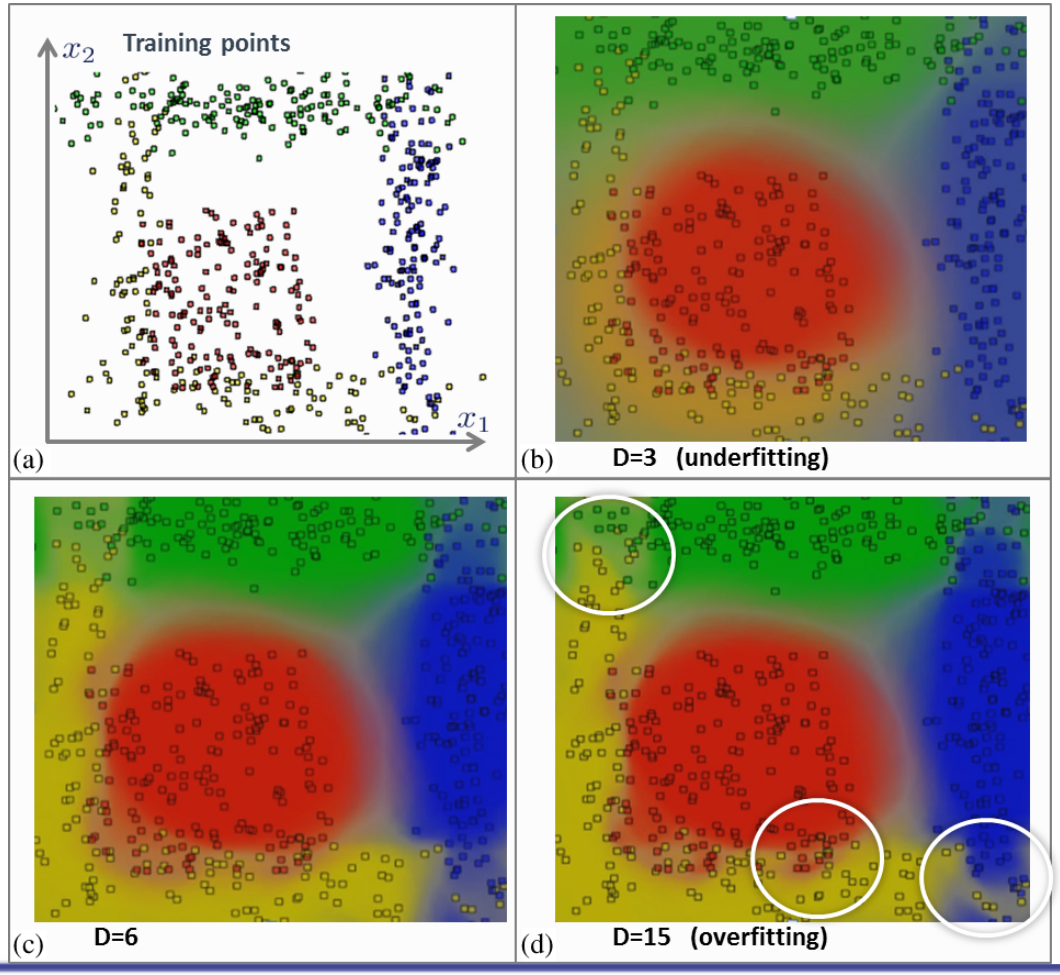
What do we gain by using a Forest?



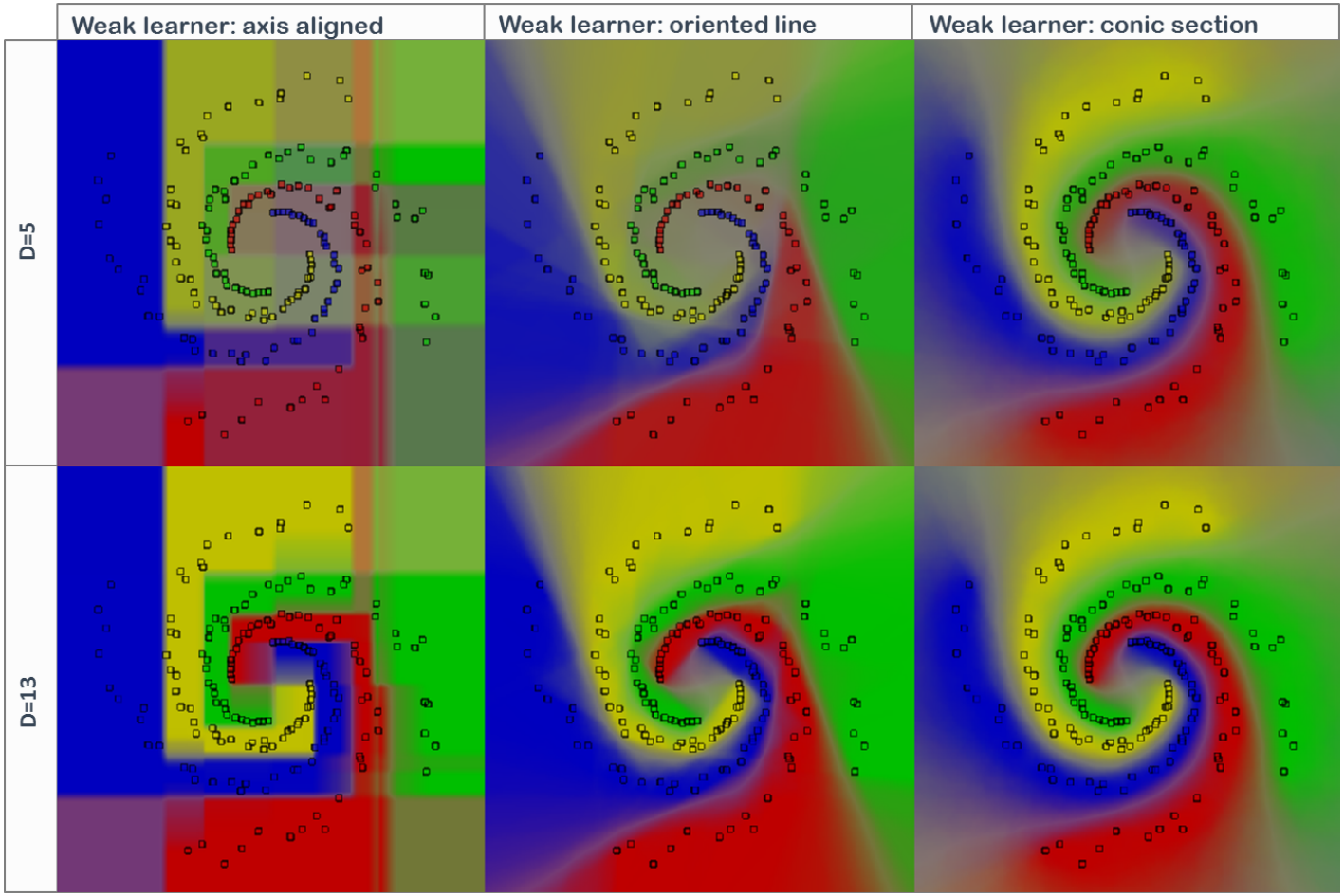
Noise Resilience and Topology Independence



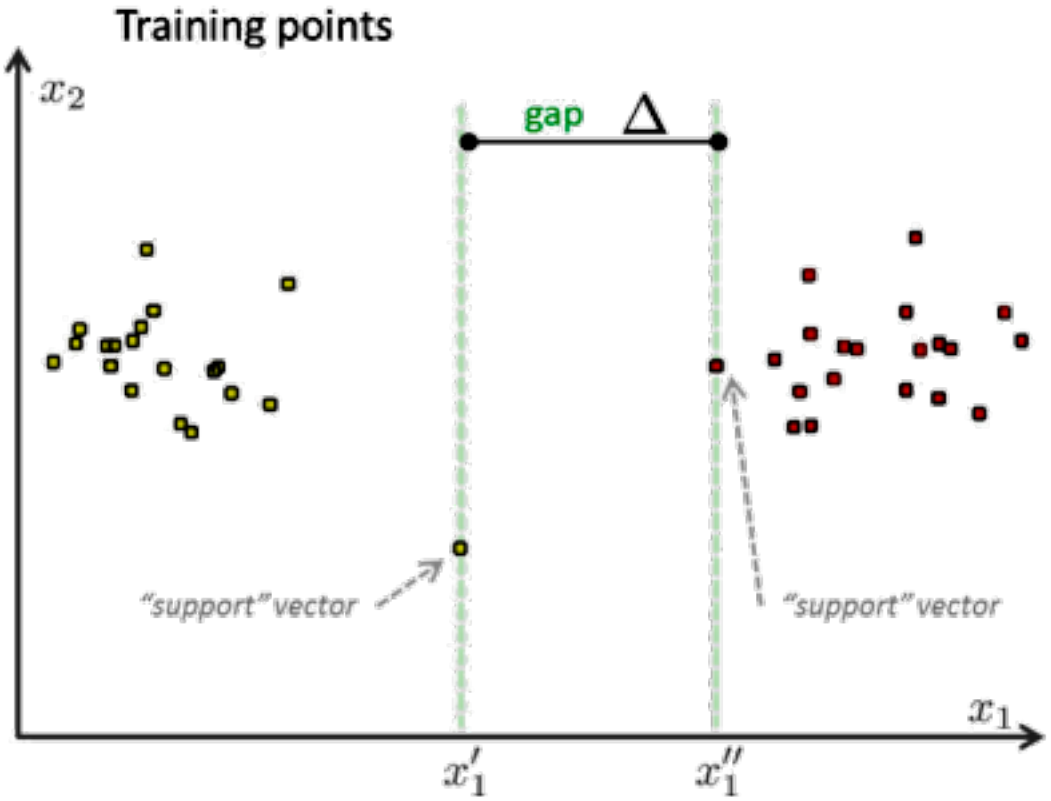
Effect of Tree Depth

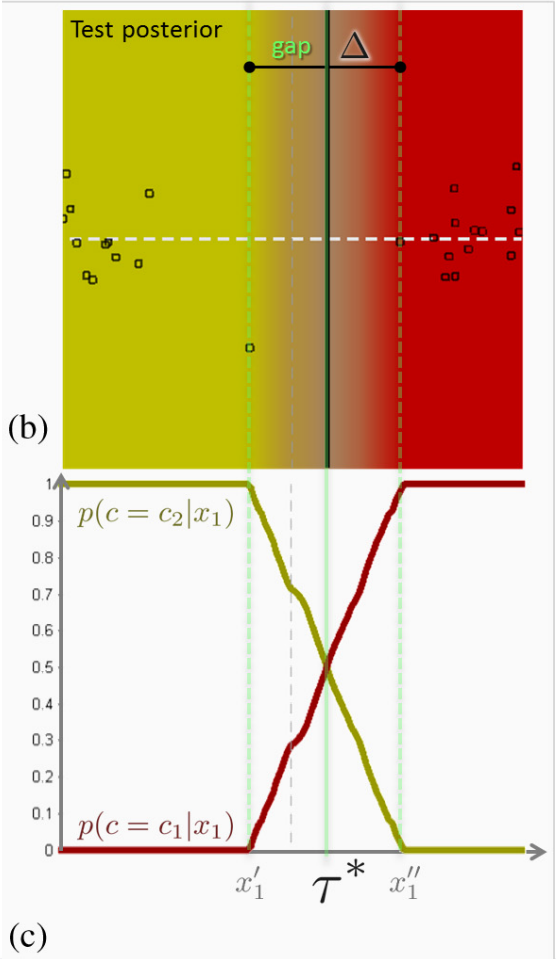


Effect of Split Function

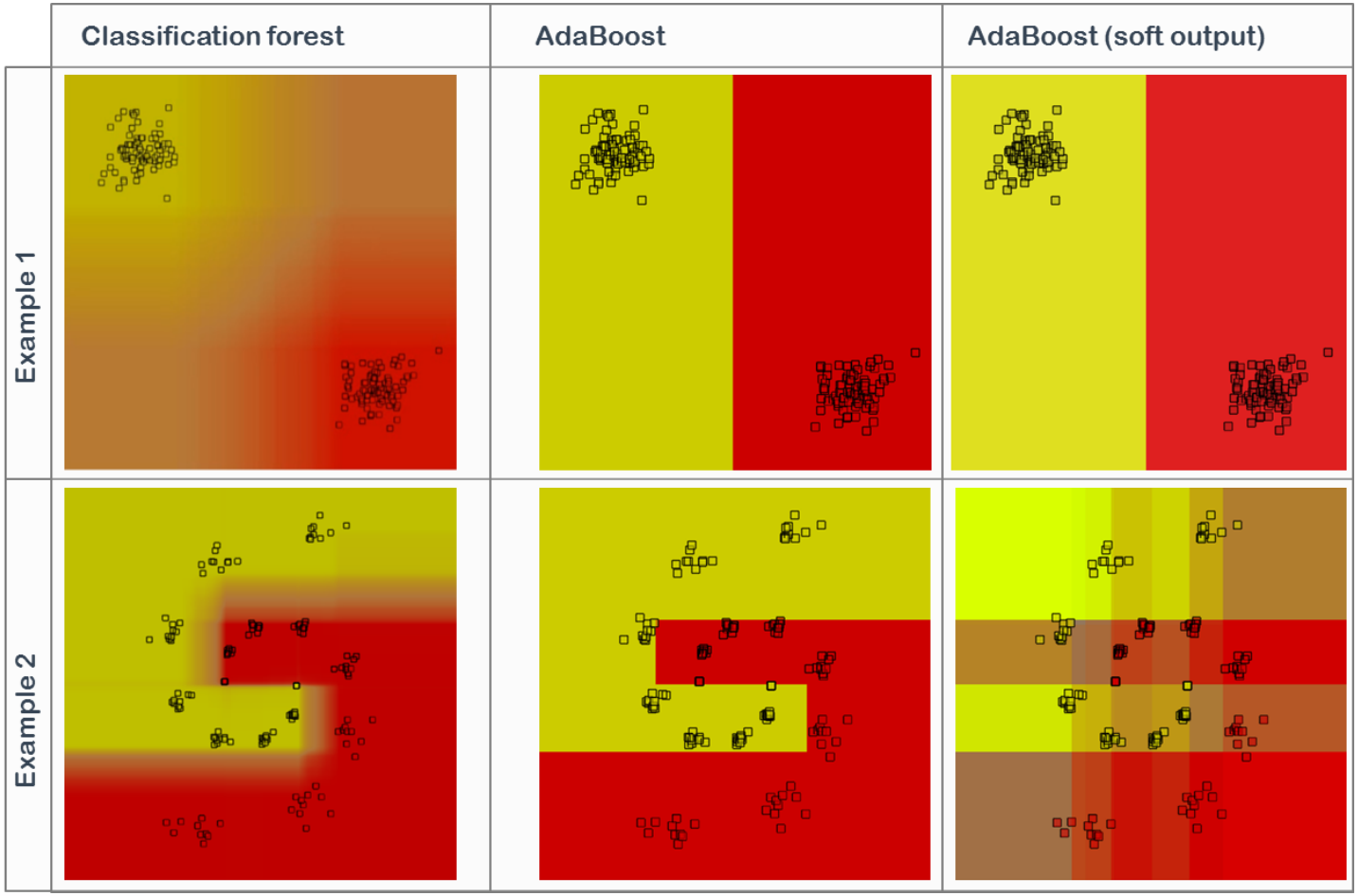


Classification Margin

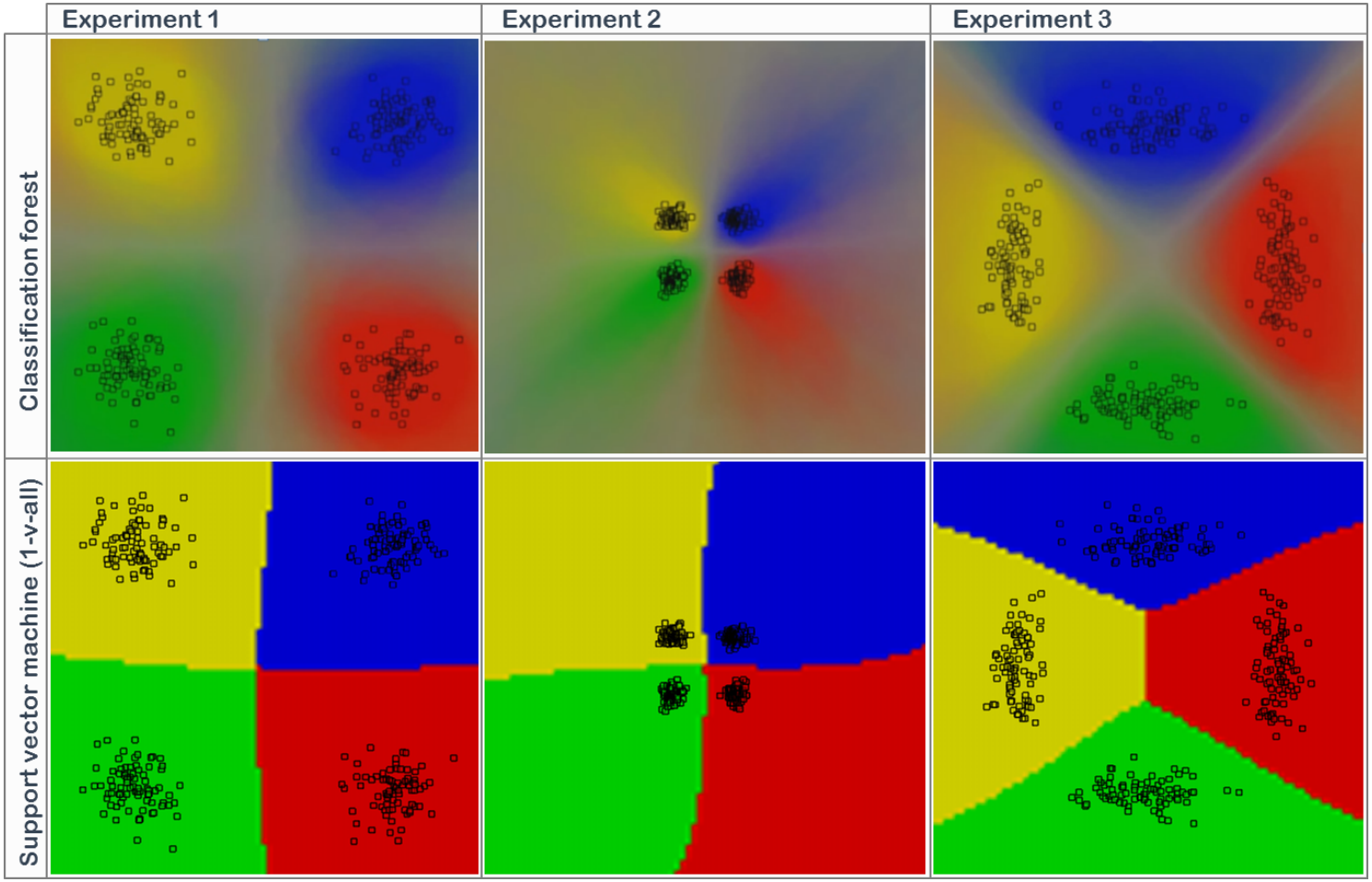




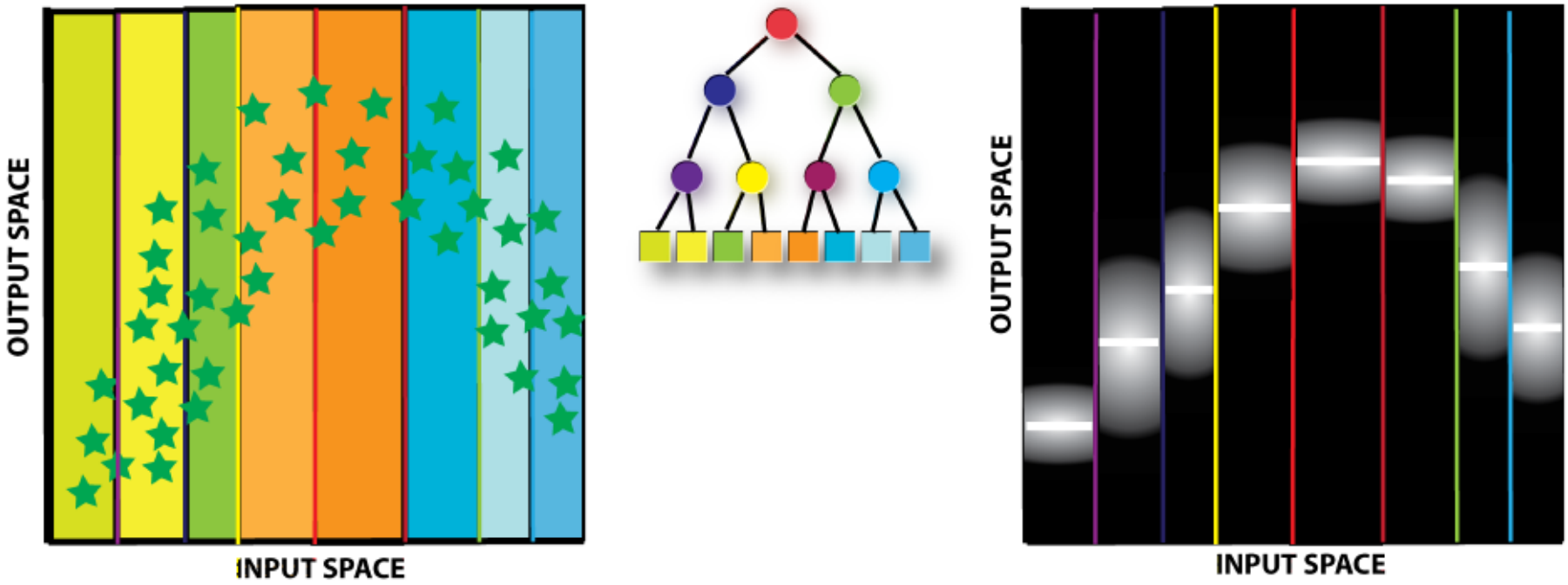
Random Forest vs. AdaBoost



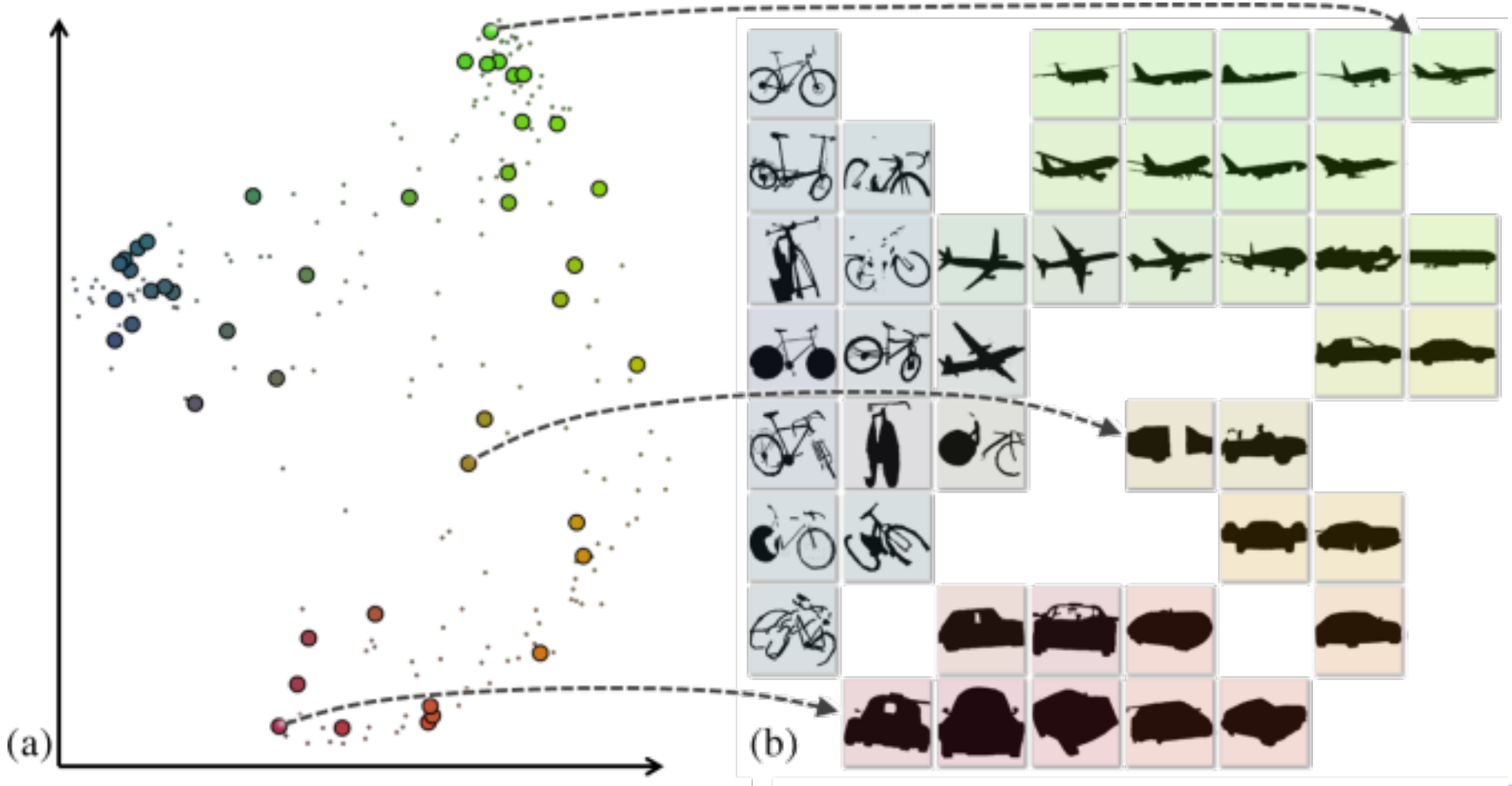
Random Forest vs. SVM



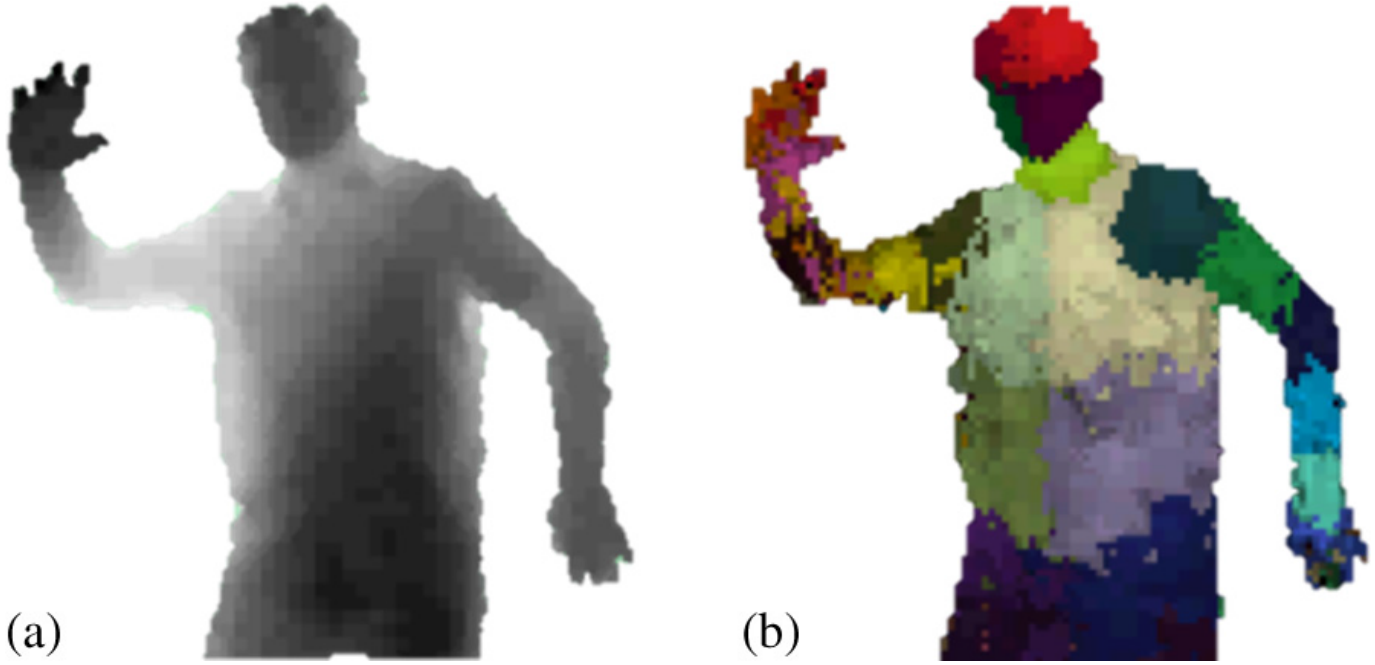
Regression Forest



Manifold Forest



Gaming – Kinect for Xbox 360



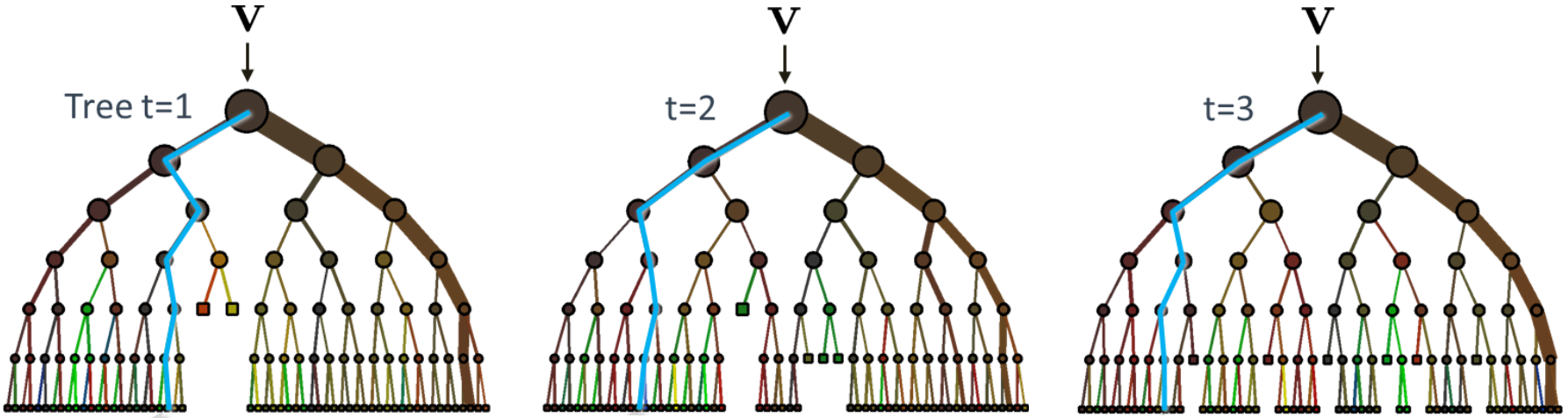
Depth map

Body part classification

J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time human pose recognition in parts from a single depth image," in *Proc. CVPR, 2011.*

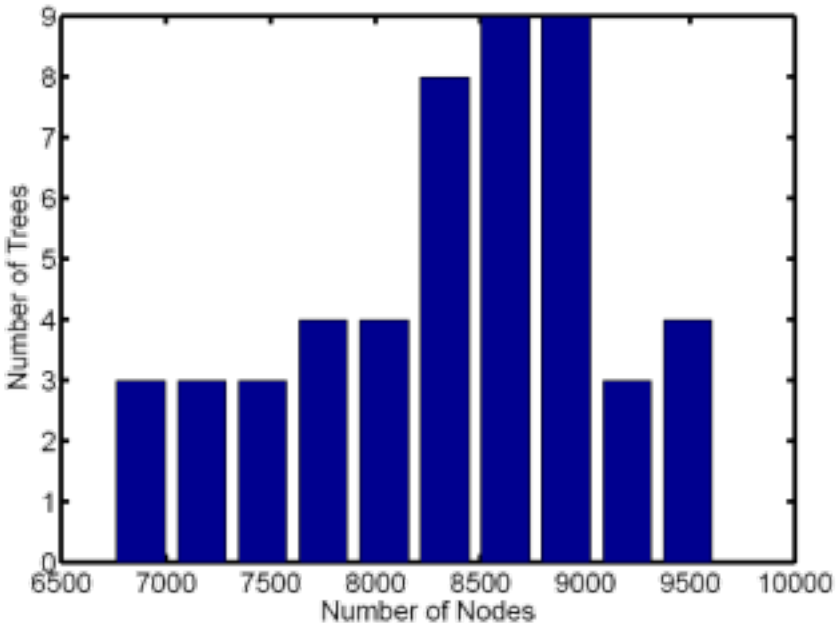
ENGINEERING DESIGN PERSPECTIVE

Understanding Computations

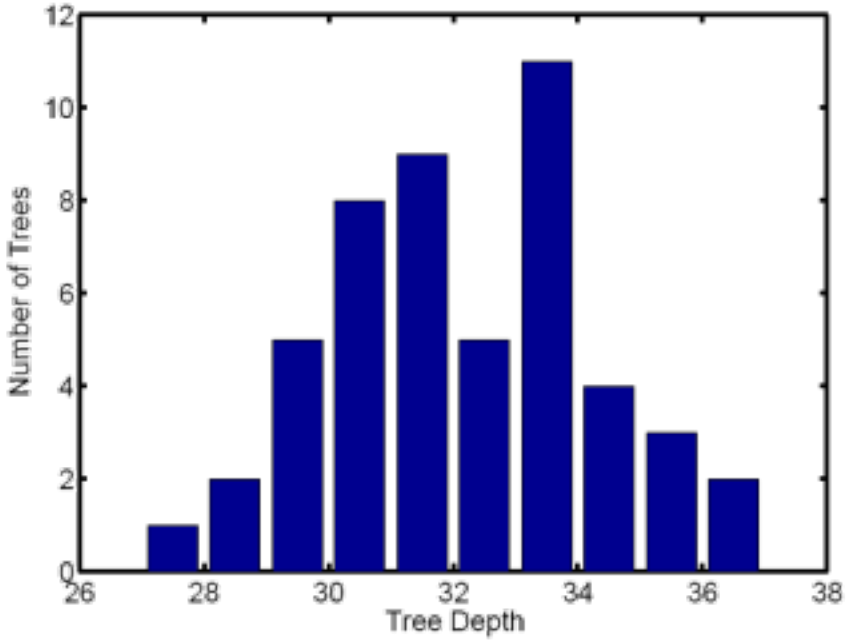


Computational Complexity

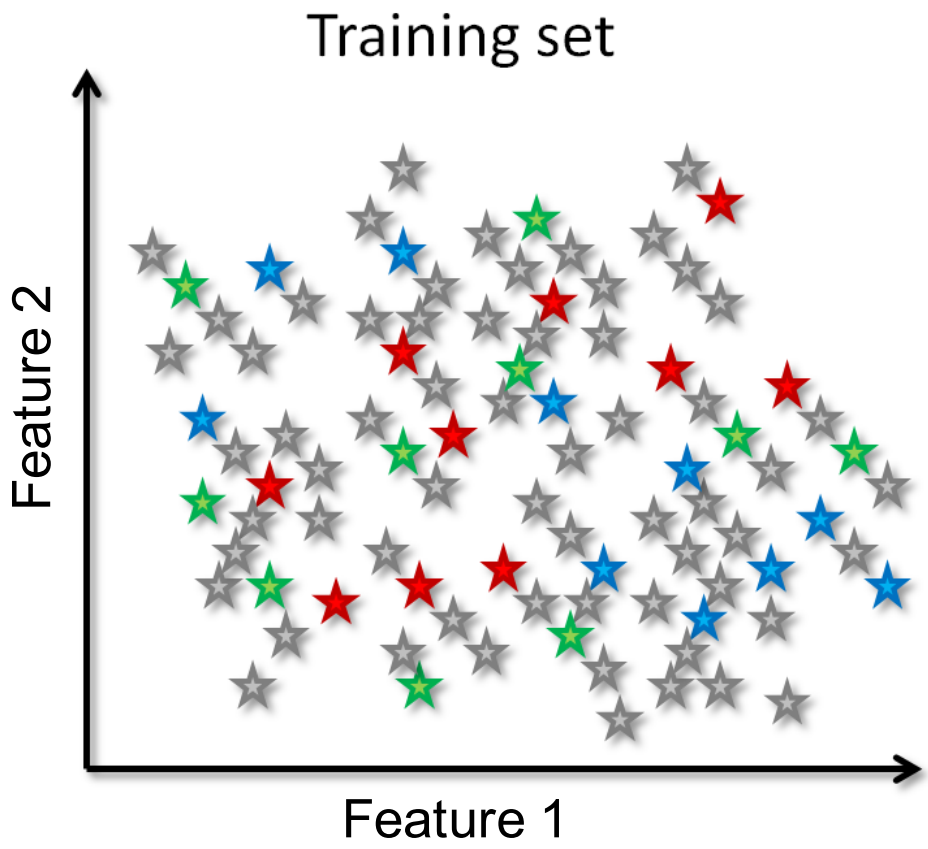
Training Complexity

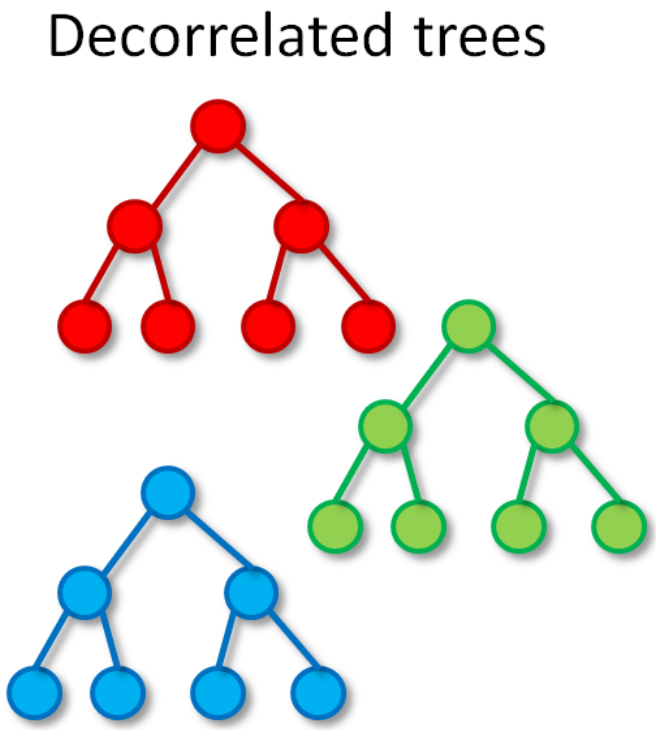


Testing Complexity

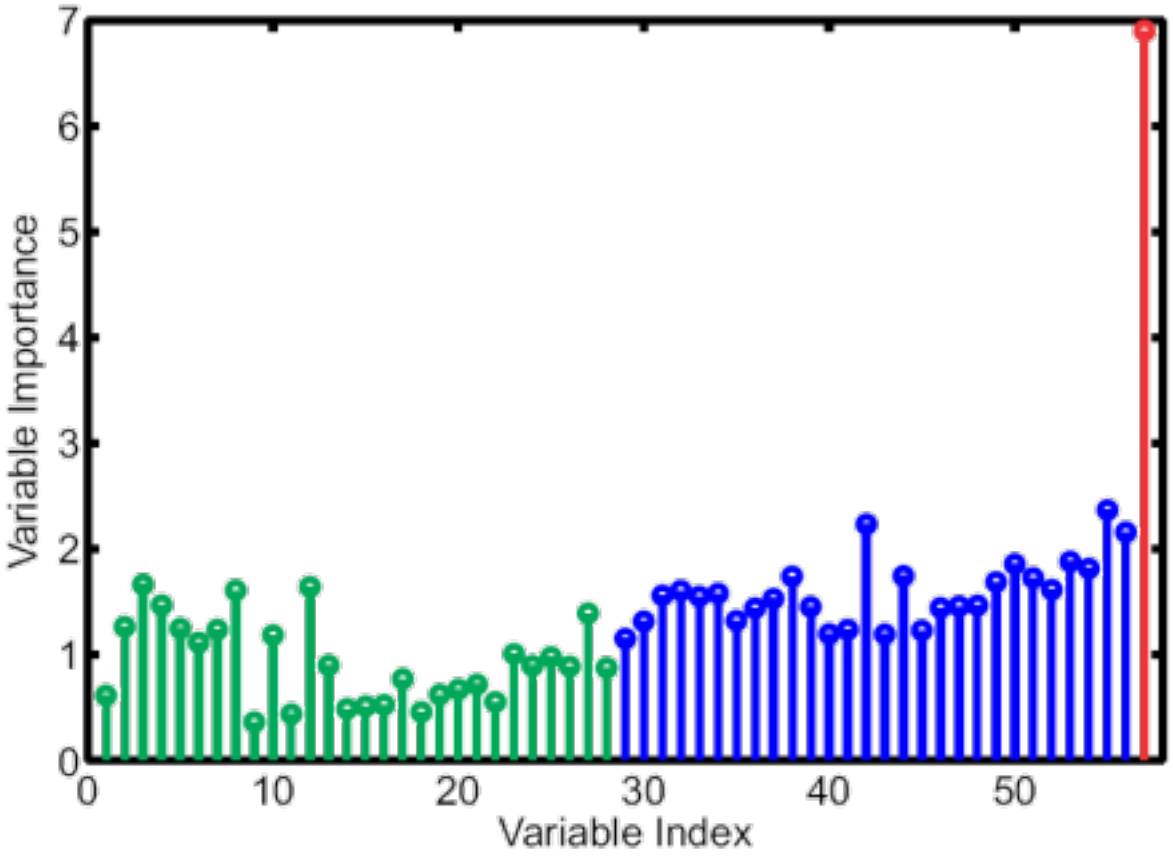


Features and their Role





Variable Importance



Genuer, R., Poggi, J.-M., Tuleau-Malot, C., (2010). Variable selection using random forests. *Pat. Recog. Letters*. **31**(14):2225-2236

Take Home Message

- Reading
  - L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Chapman and Hall/CRC, 1984.
  - L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
  - A. Criminisi and J. Shotton, *Decision Forests for Computer Vision and Medical Image Analysis*, Springer, 2013.
- Toolboxes and Packages
  - randomForest in R
  - TreeBagger in Matlab
  - sklearn.ensemble.RandomForestClassifier in Python-Scikit-learn
- Conferences
  - Int. Conf. Comp. Vis. (ICCV)
  - Eur. Conf. Comp. Vis. (ECCV)
  - Asian Conf. Comp. Vis. (ACCV)
  - Comp. Vis. Patt. Recog. (CVPR)
  - Med. Image Comp., Comp. Assist. Interv. (MICCAI)