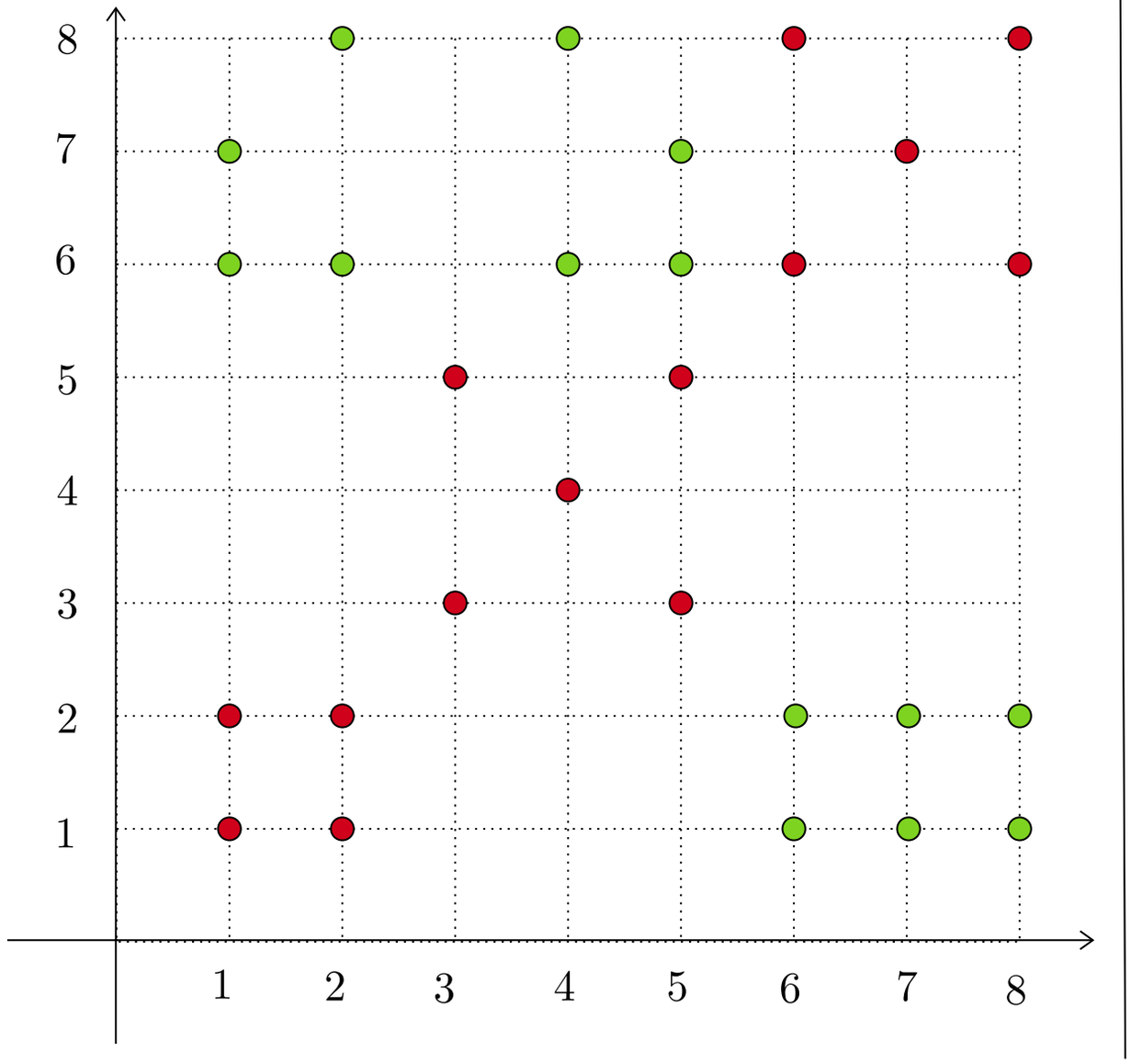
Decision Trees-2

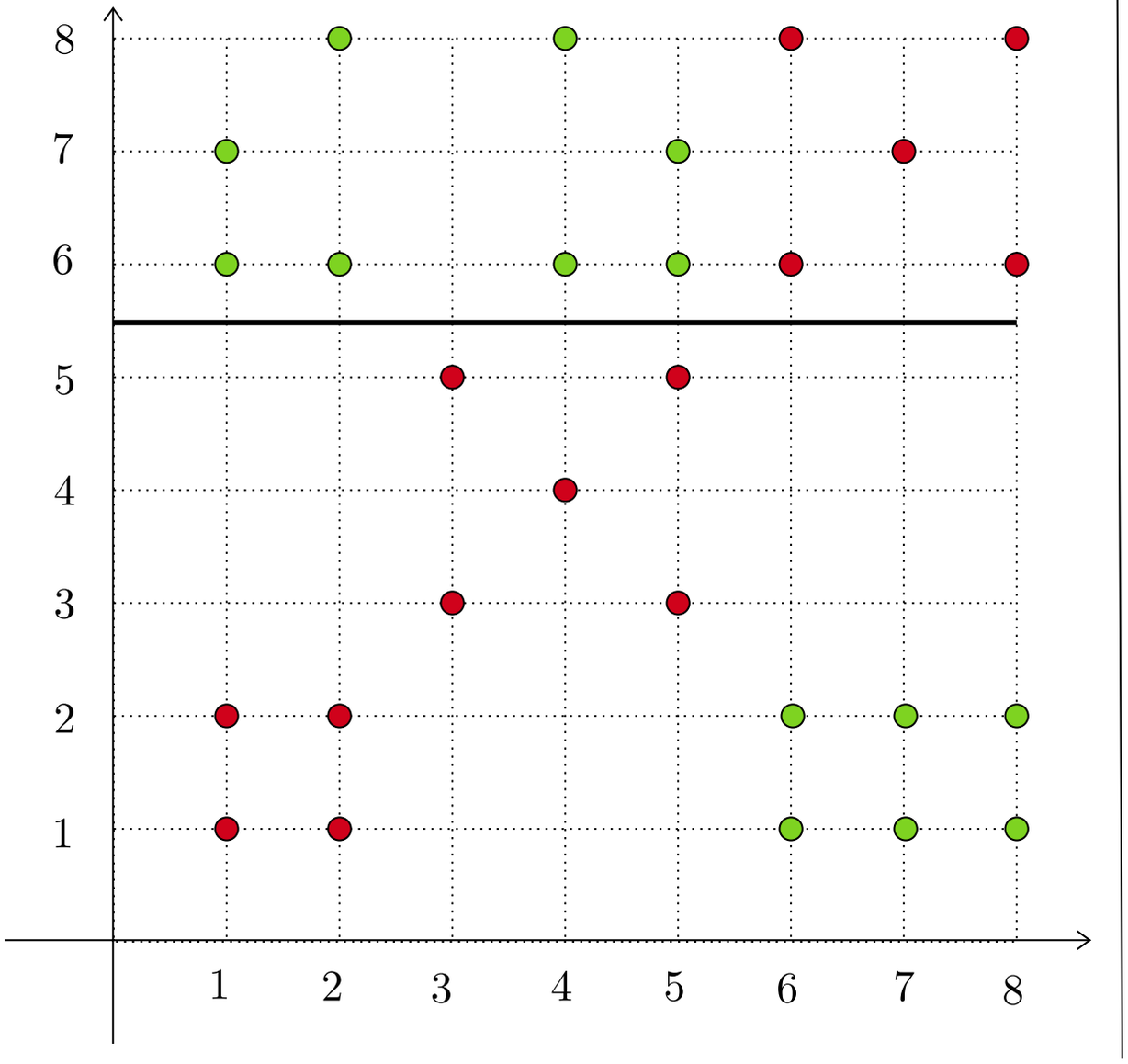
Machine Learning Techniques

Karthik Thiagarajan

Growing a Tree

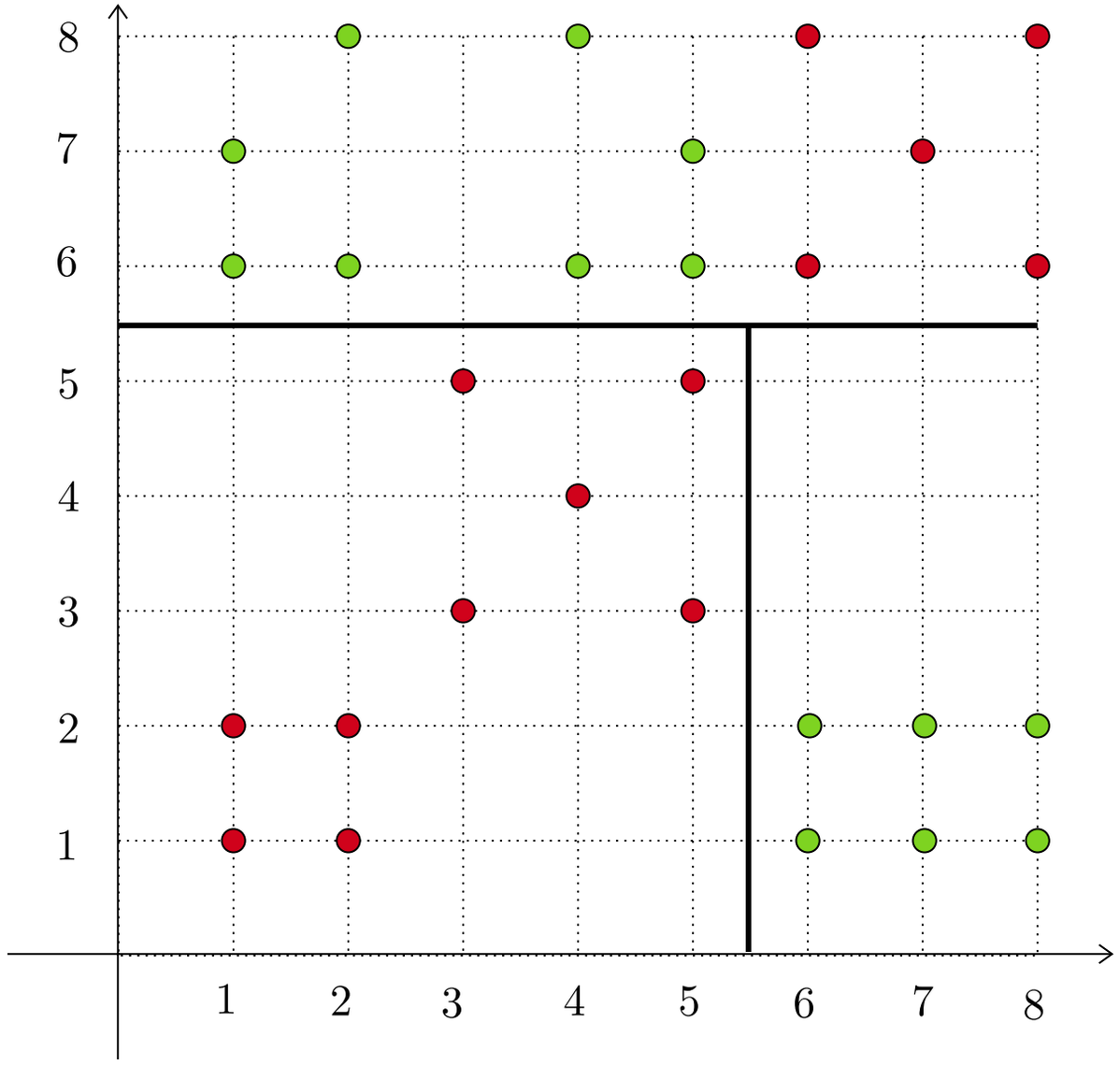


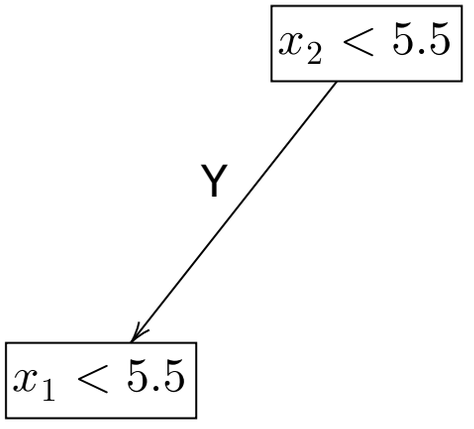
Growing a Tree



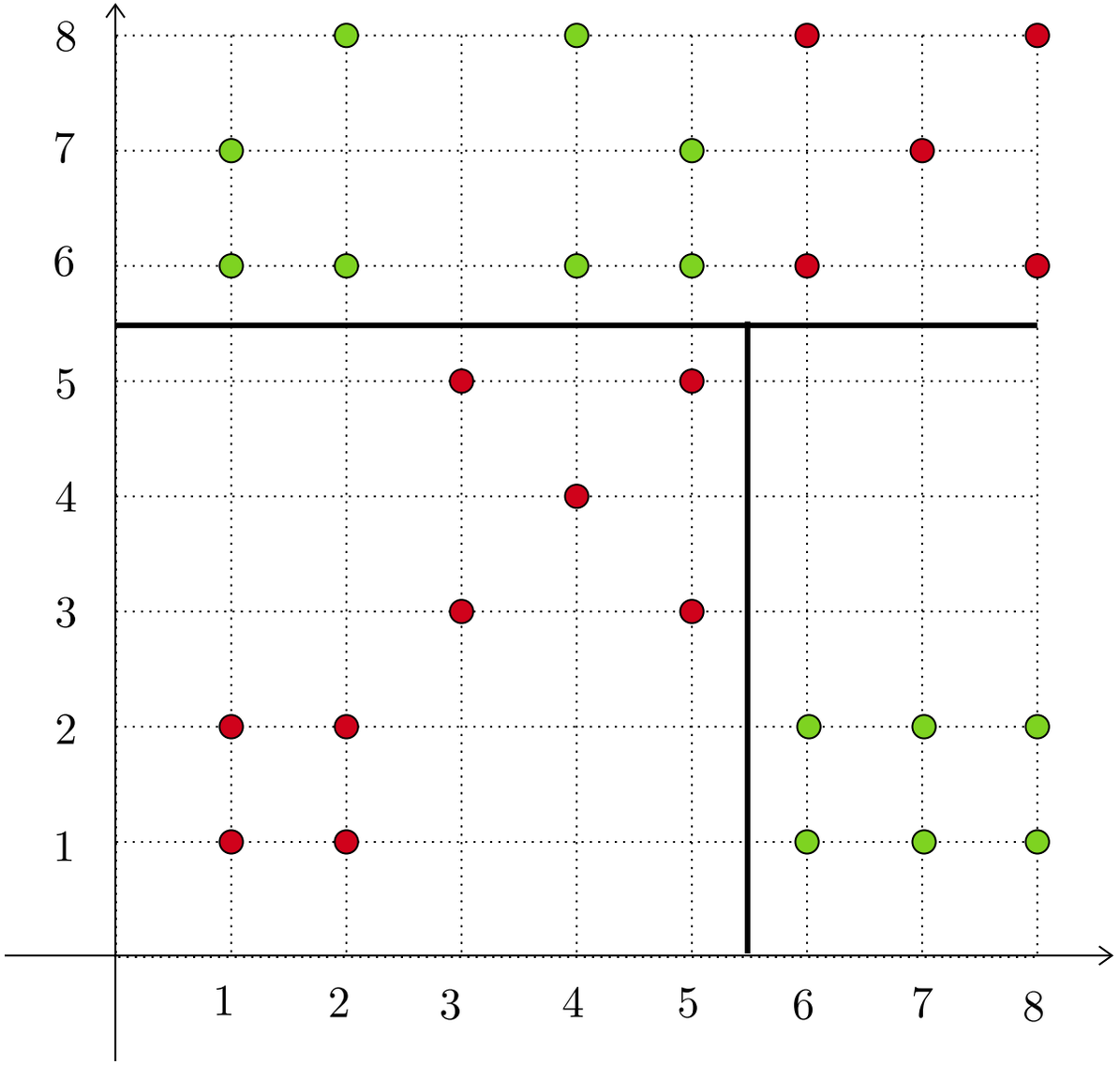
$x_2 < 5.5$

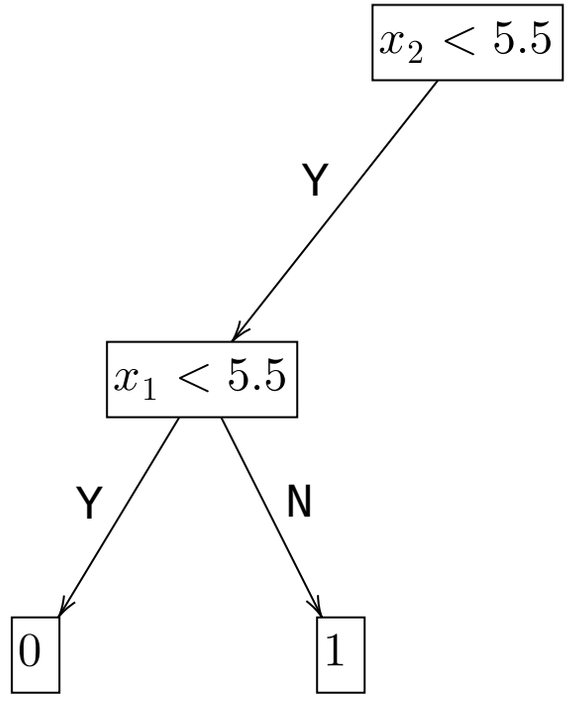
Growing a Tree



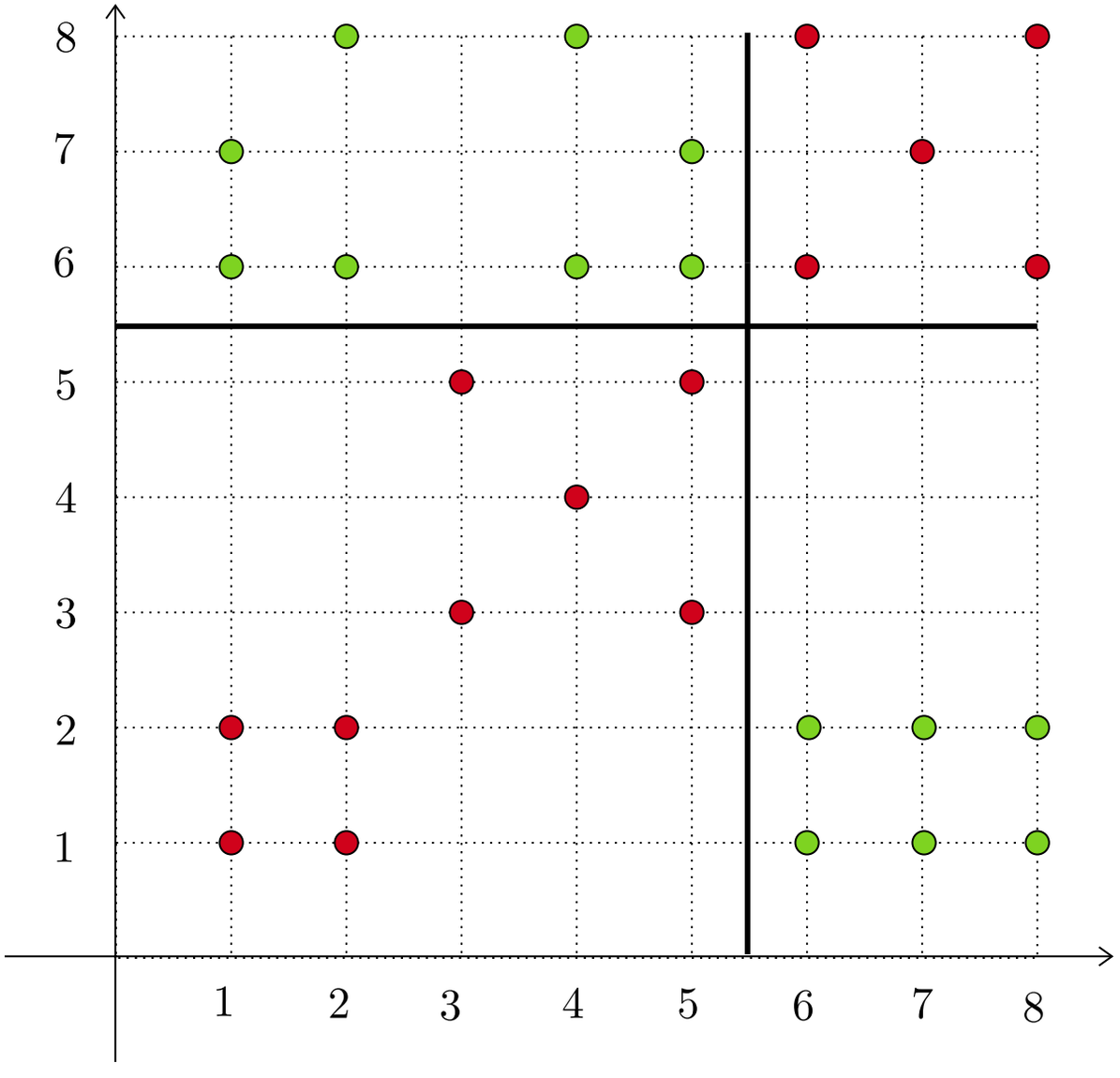


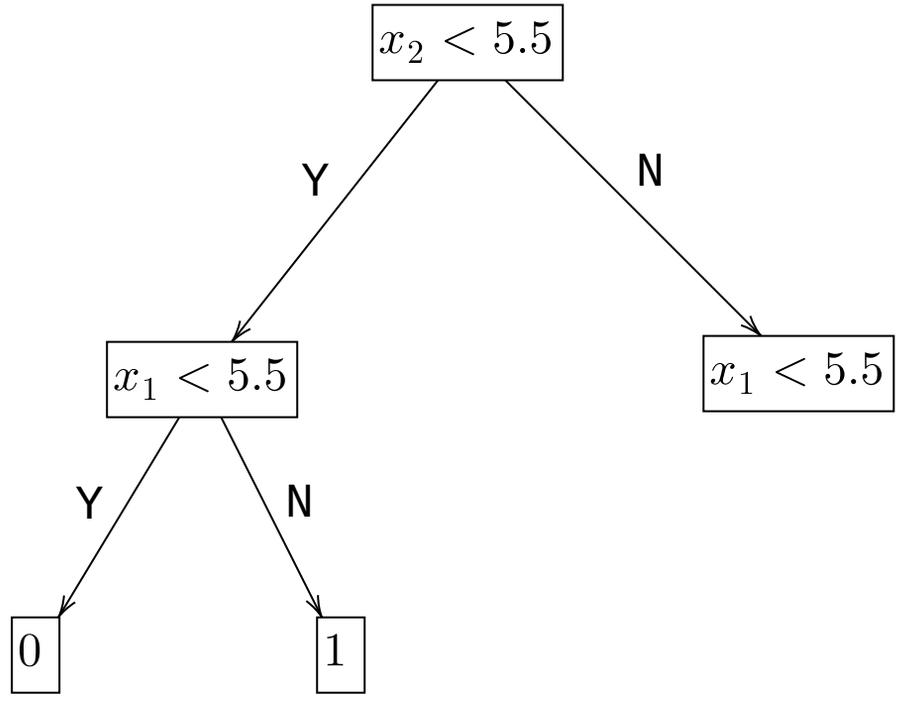
Growing a Tree



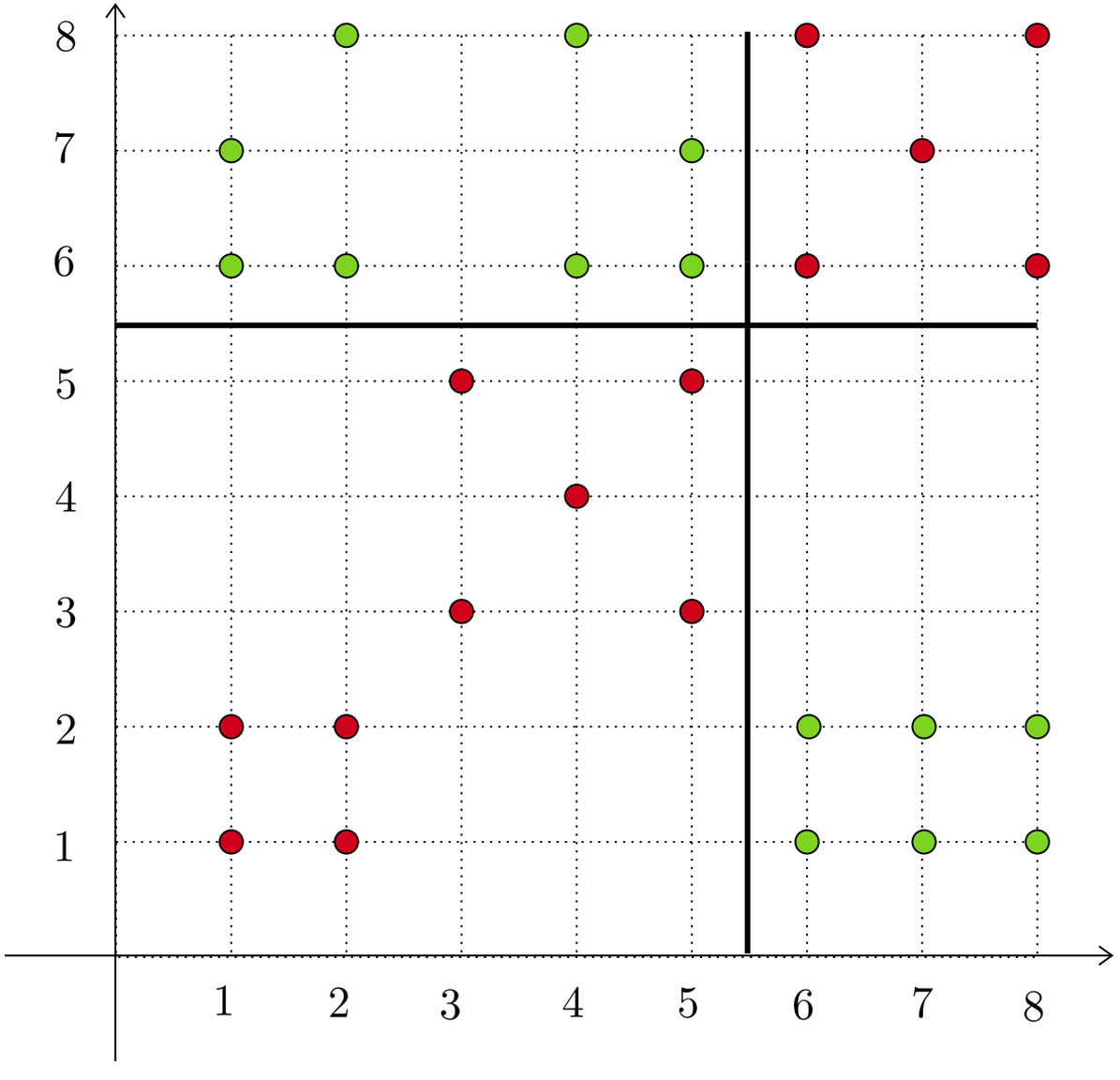


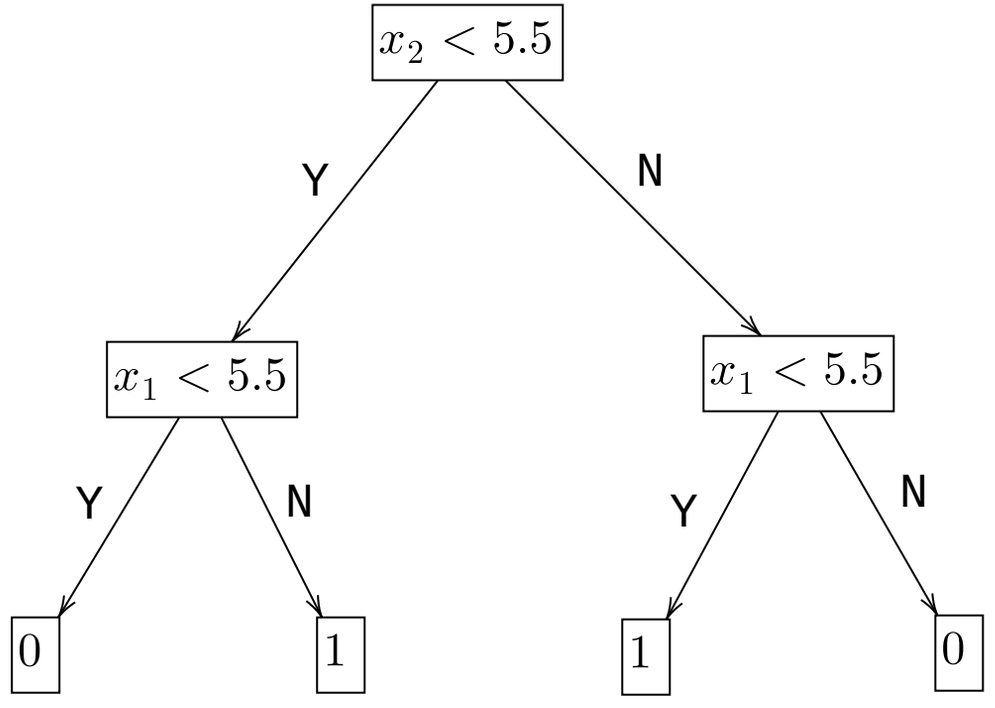
Growing a Tree



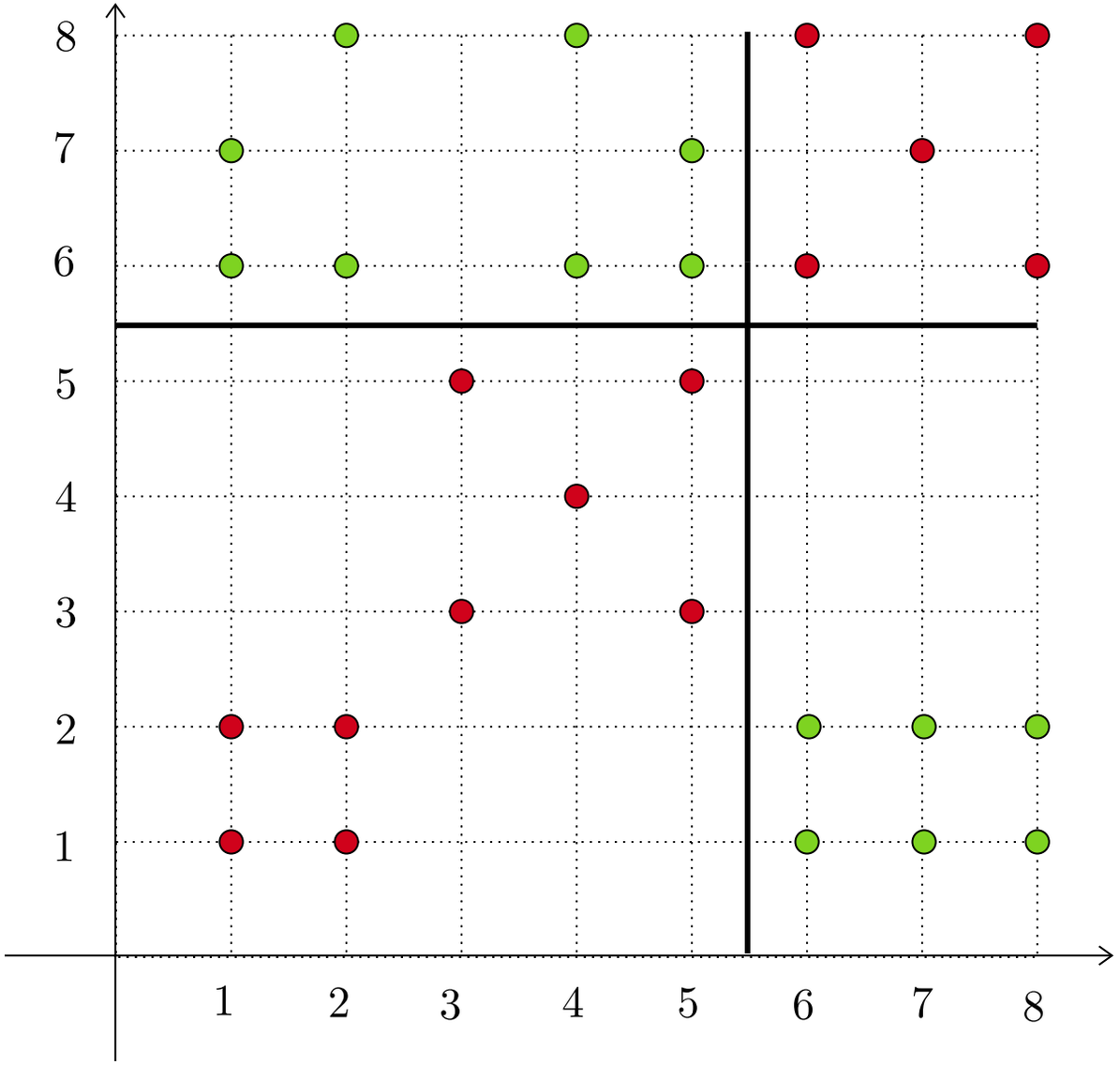


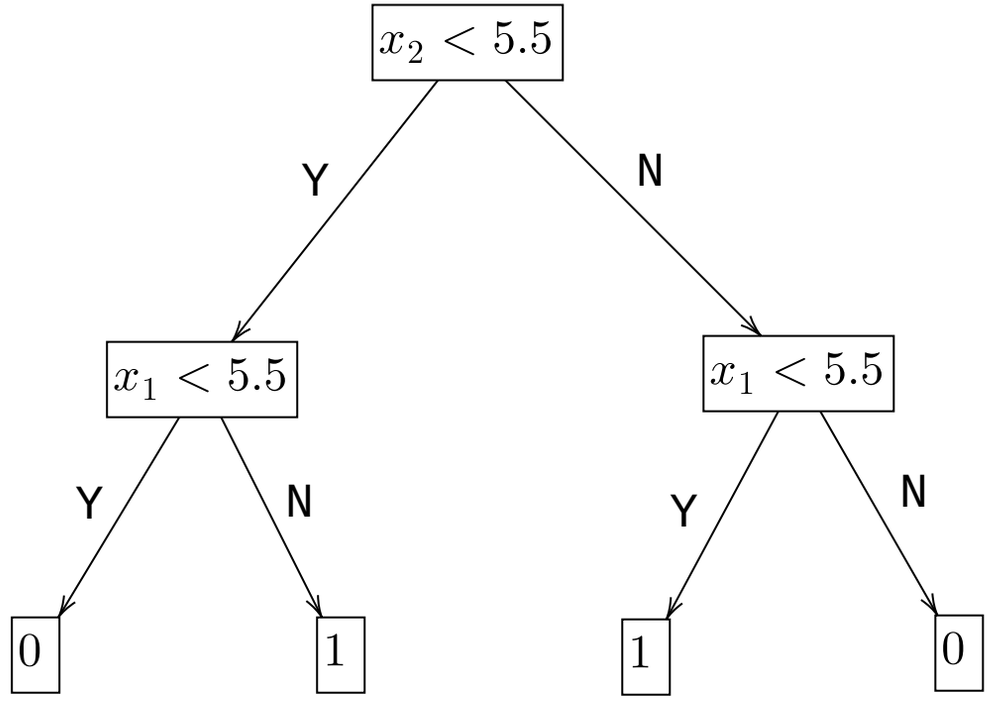
Growing a Tree

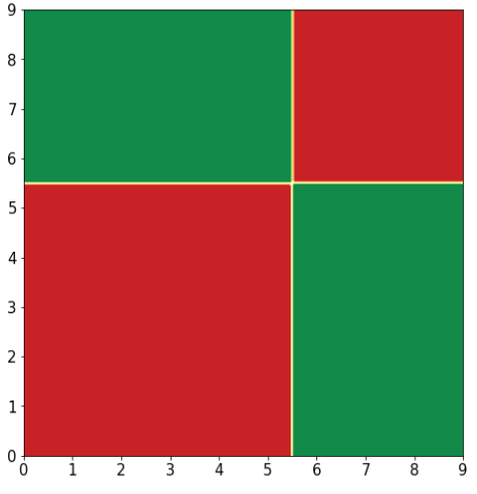




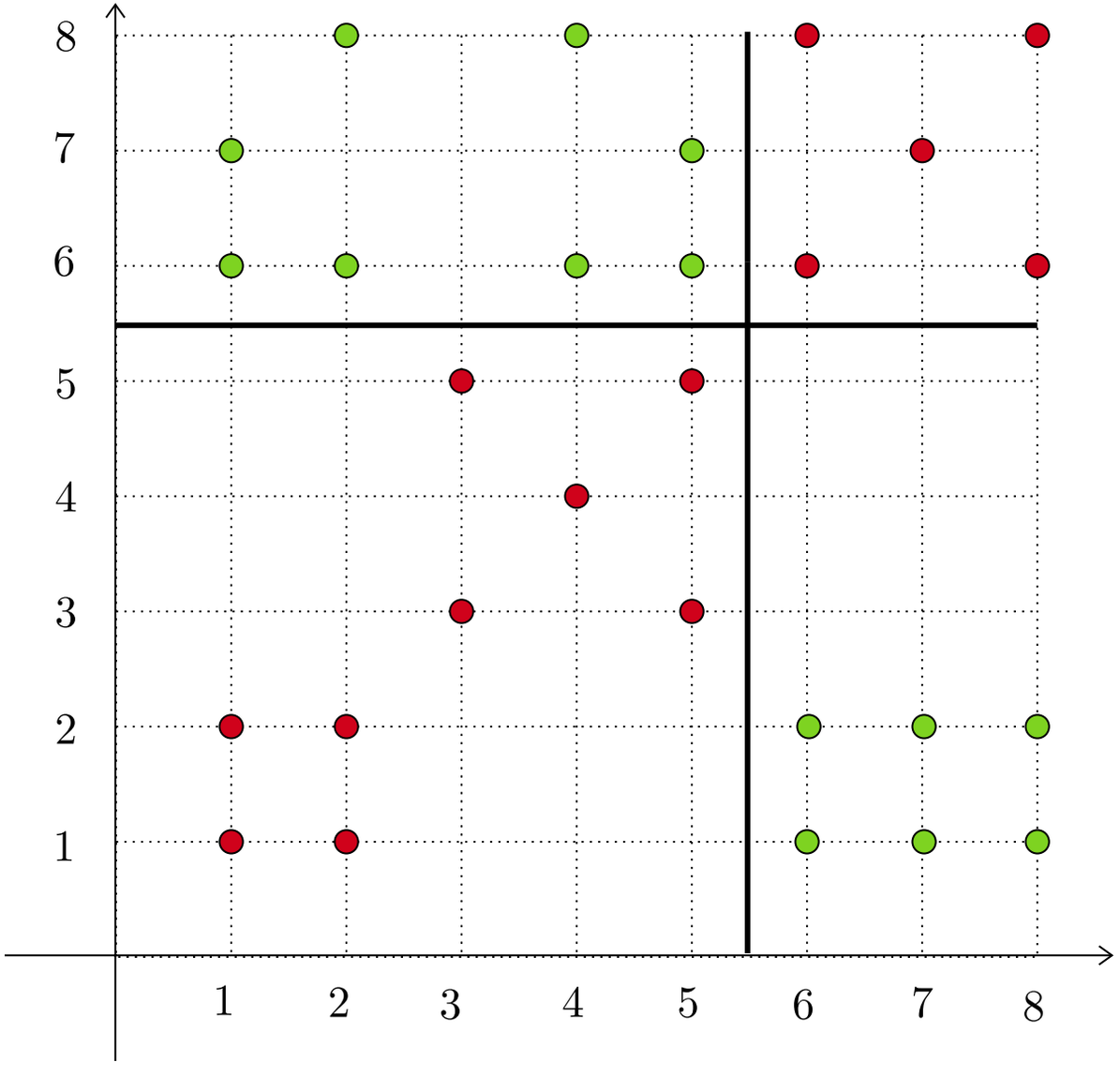
Growing a Tree

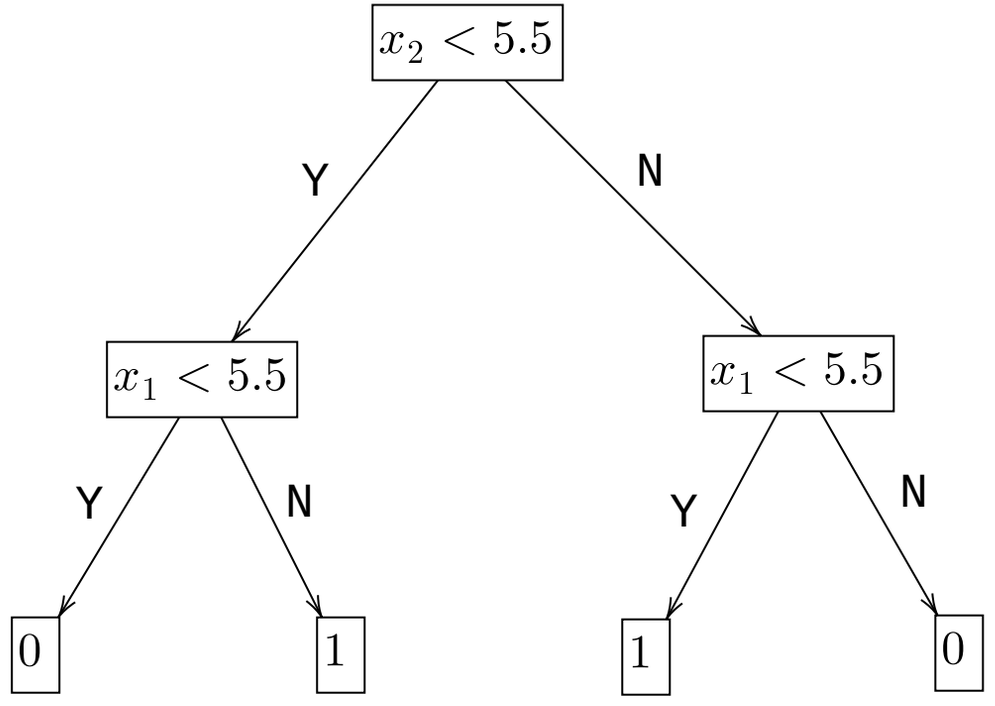


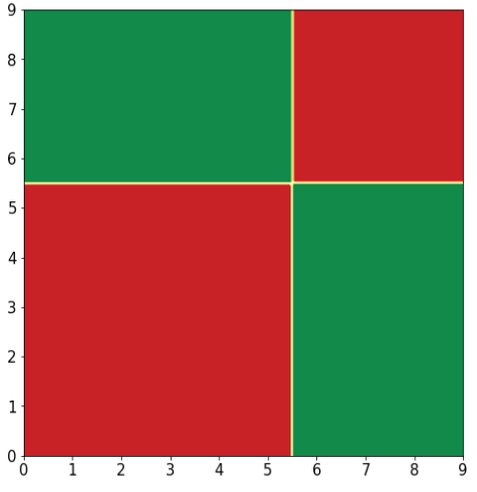




Growing a Tree

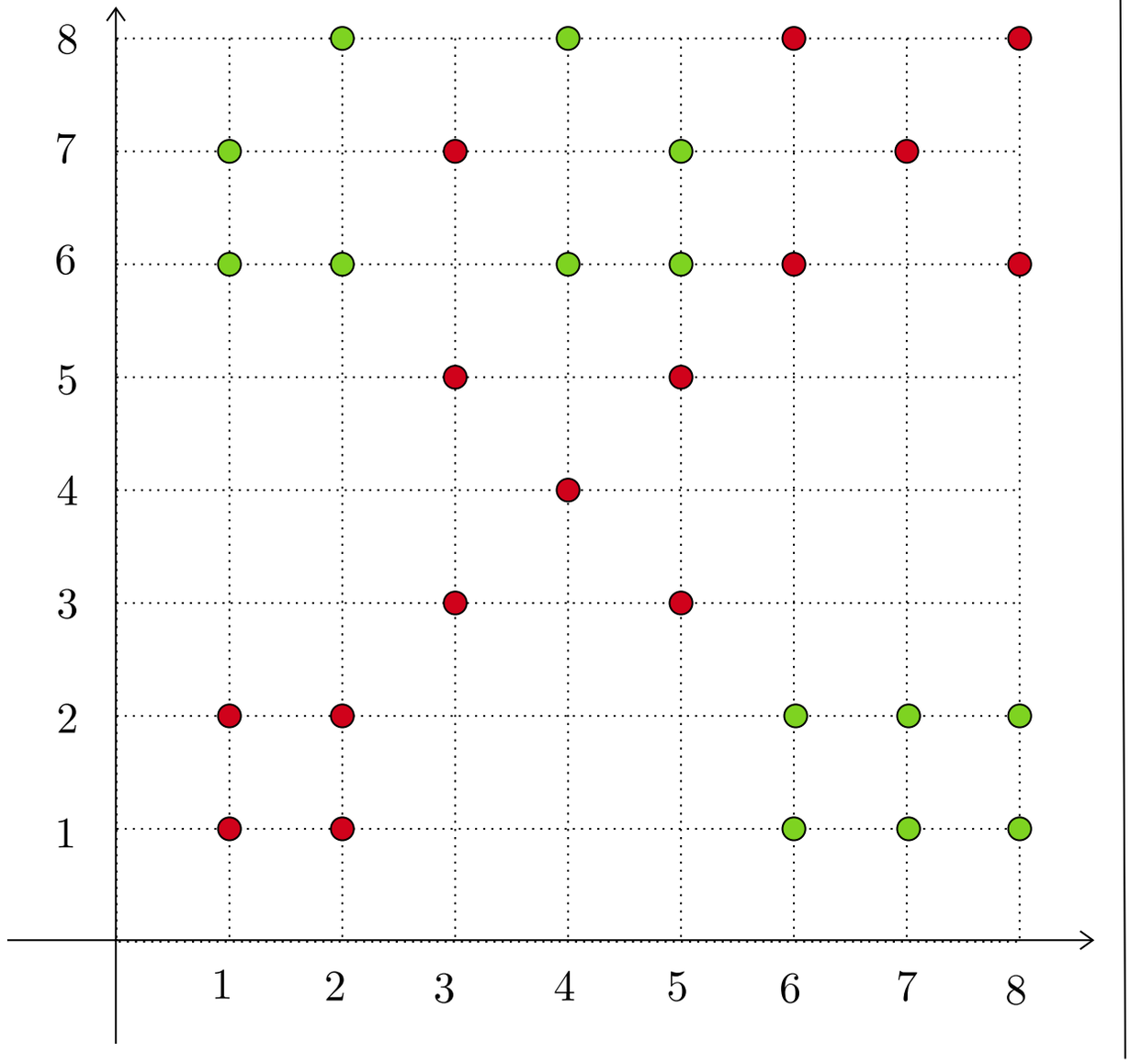




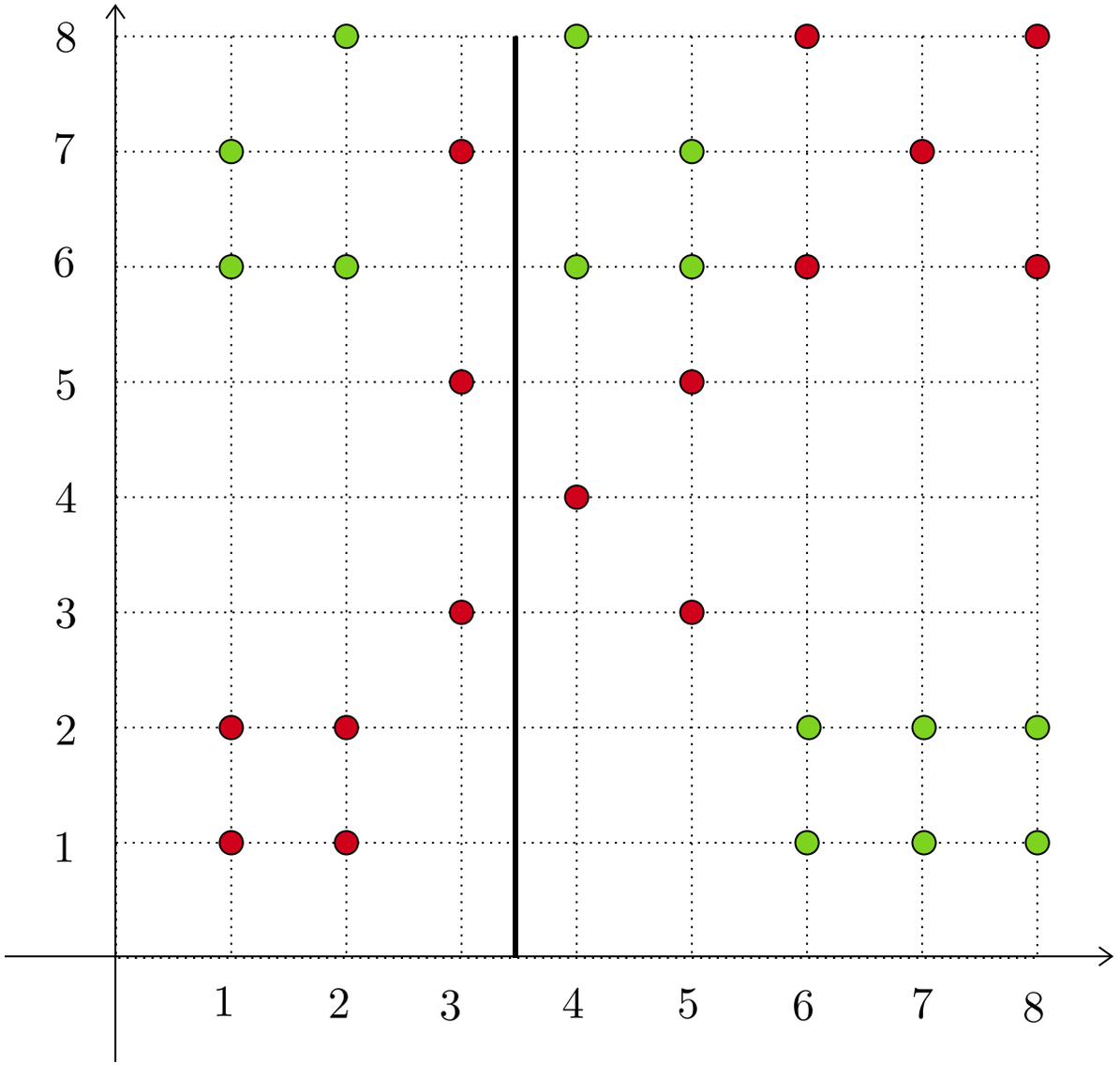


- 4 regions
- 4 leaves

Growing a Tree

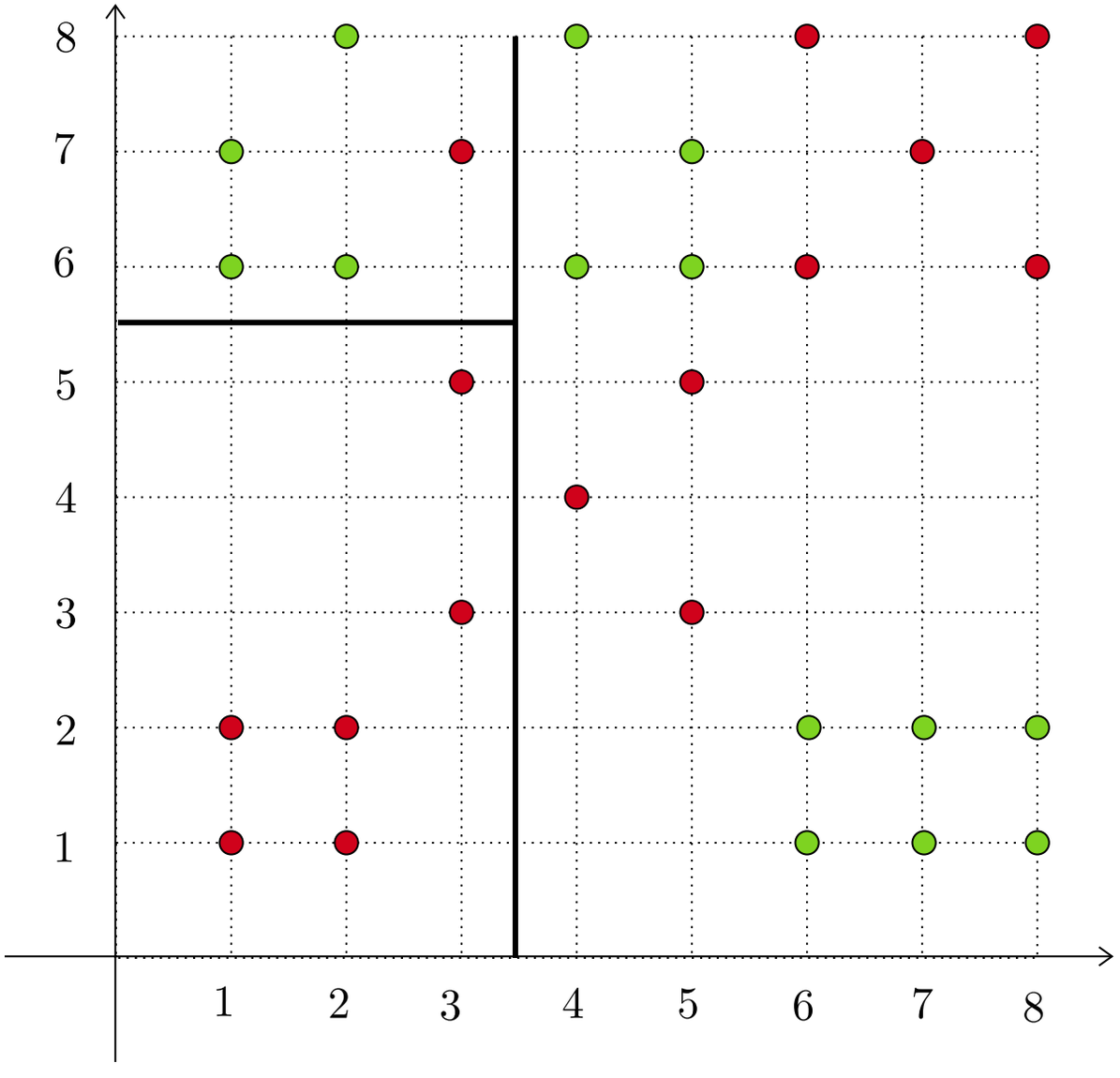


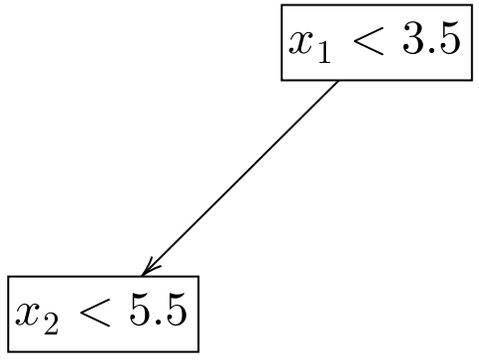
Growing a Tree



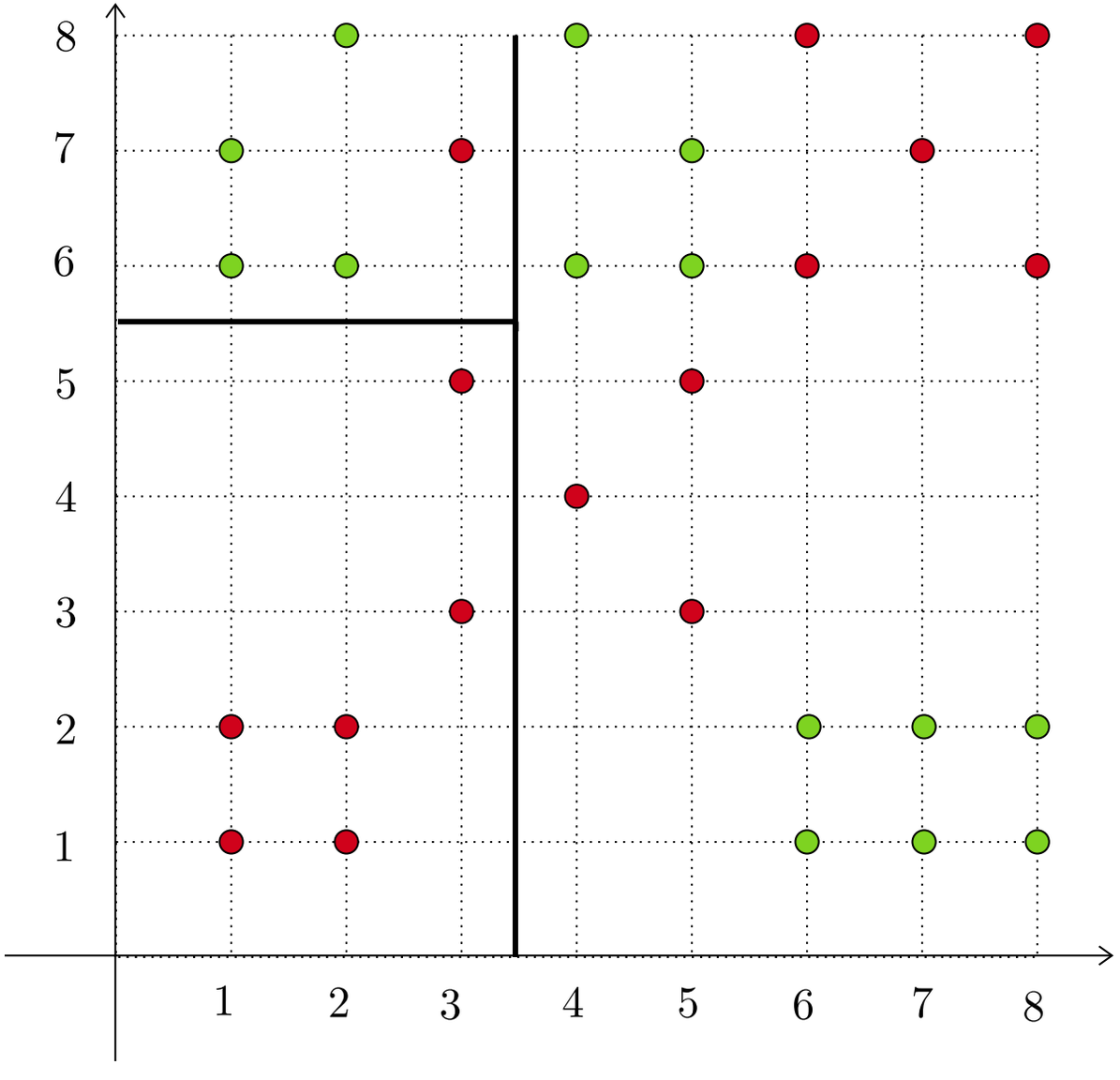
$x_1 < 3.5$

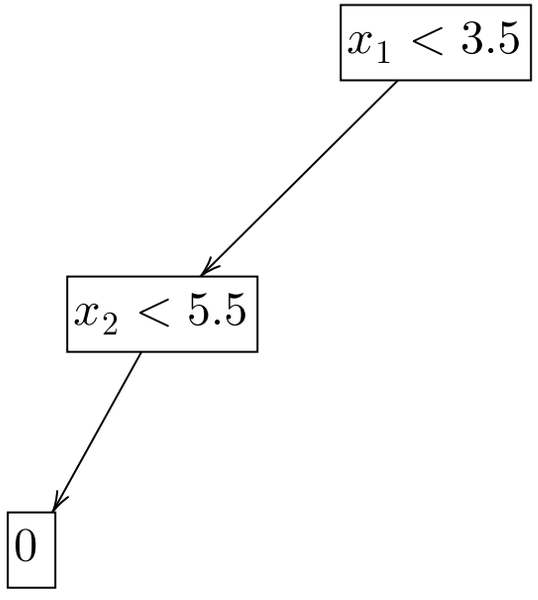
Growing a Tree



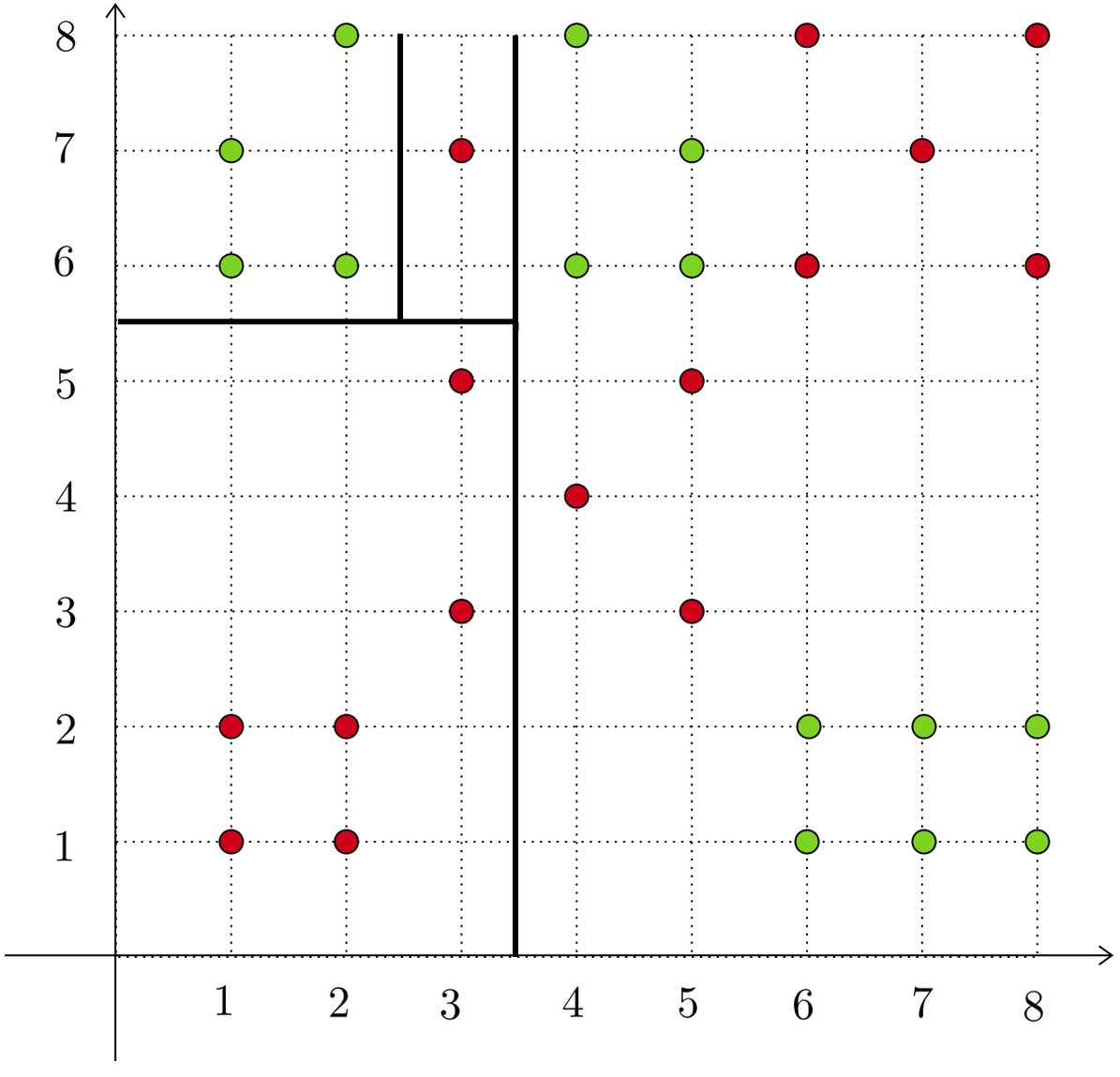


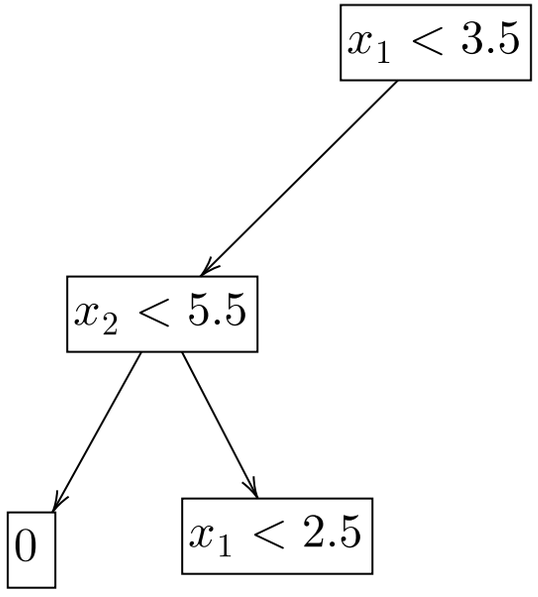
Growing a Tree



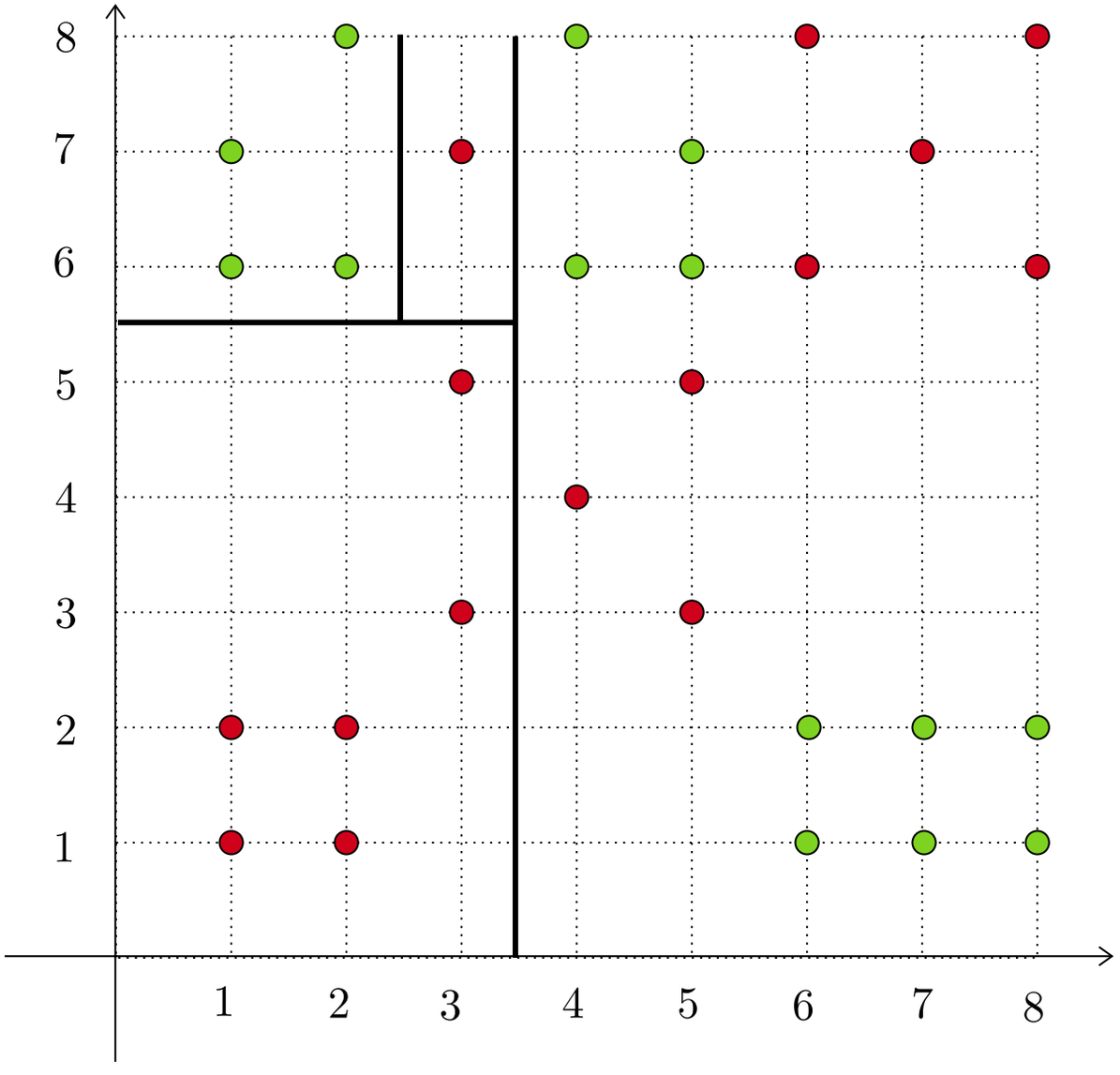


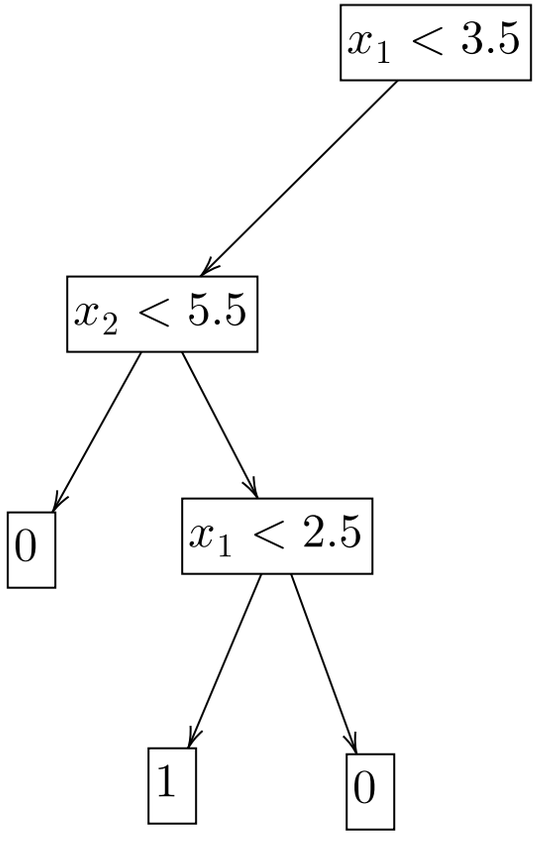
Growing a Tree



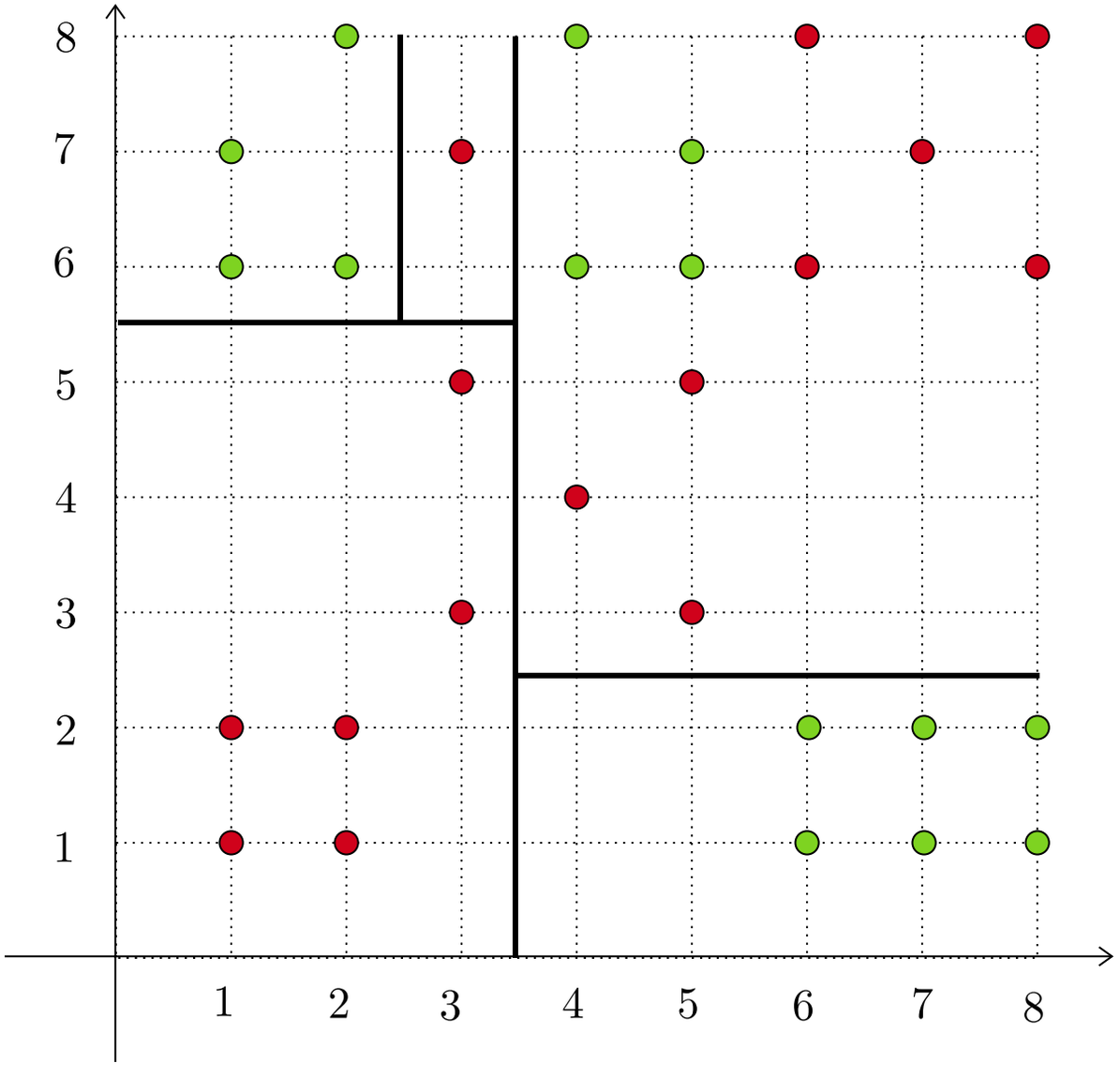


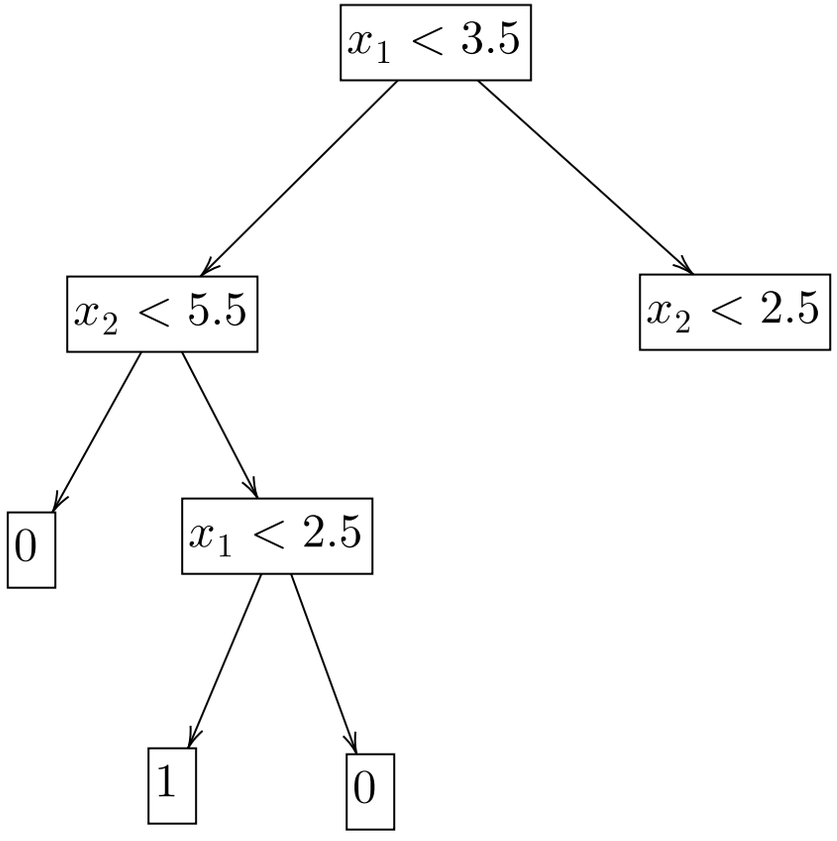
Growing a Tree



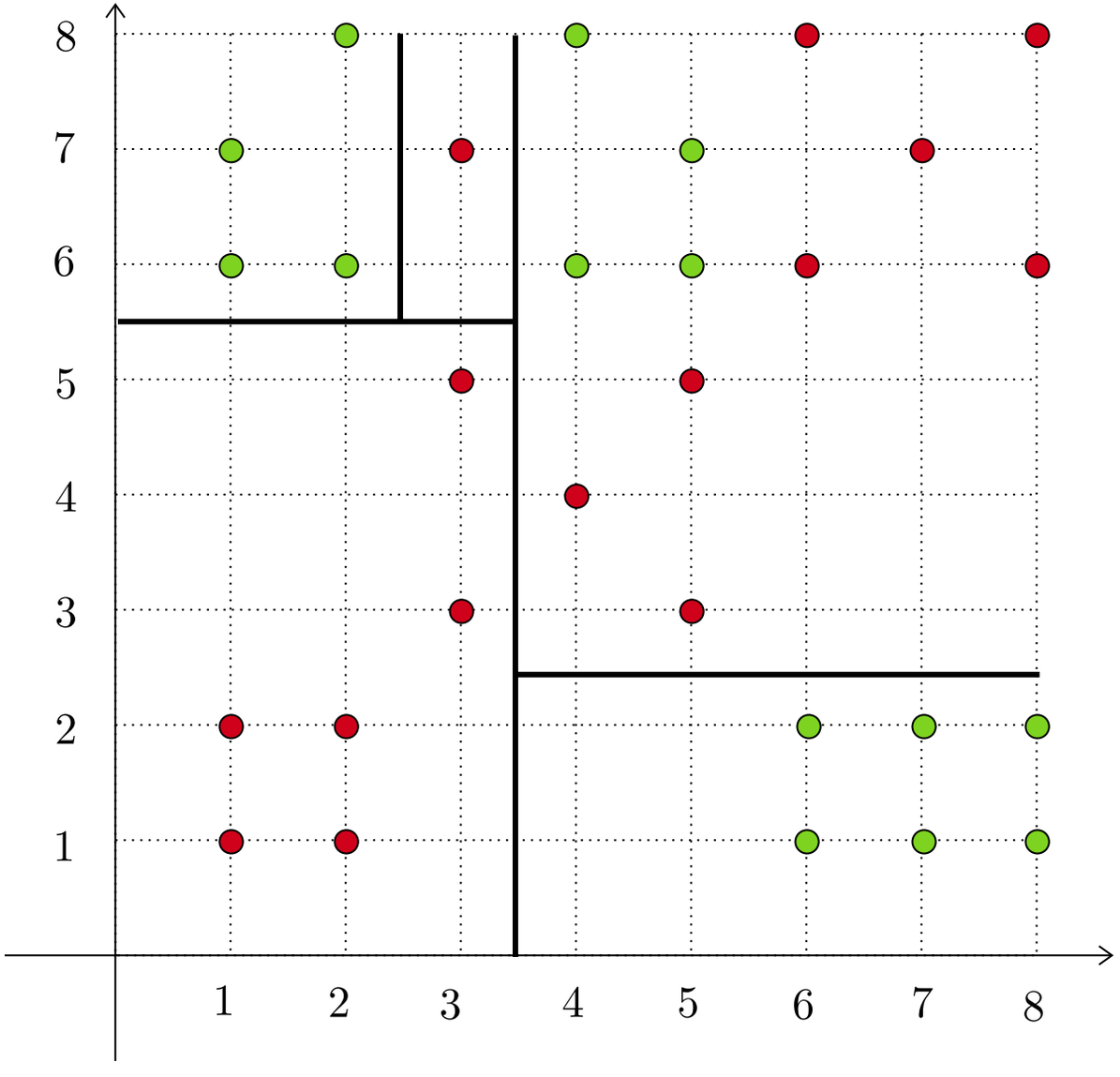


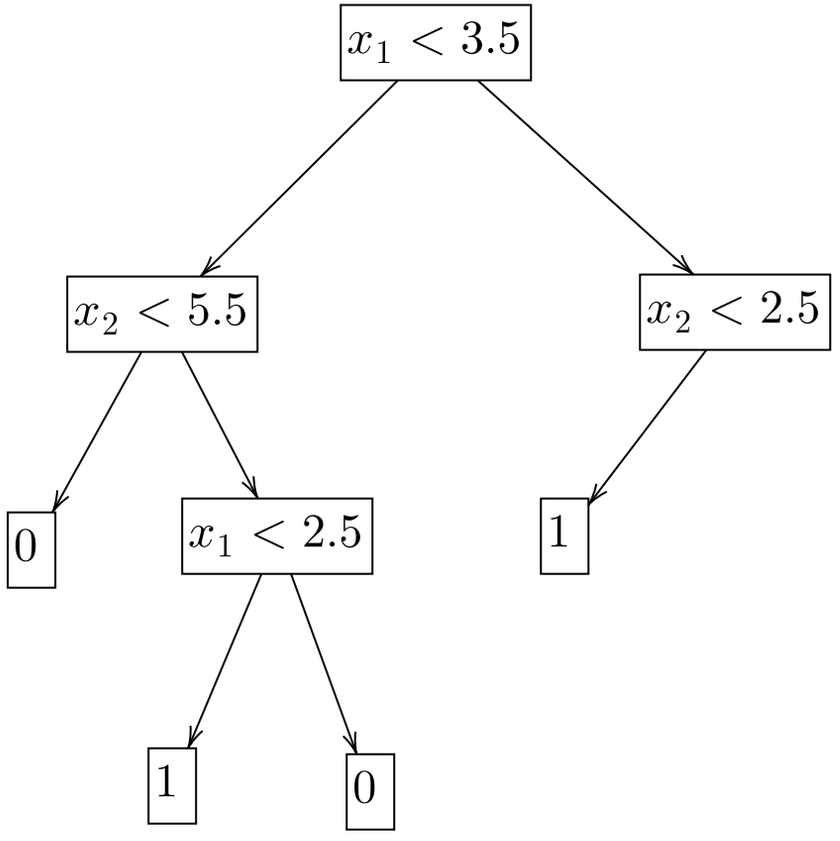
Growing a Tree



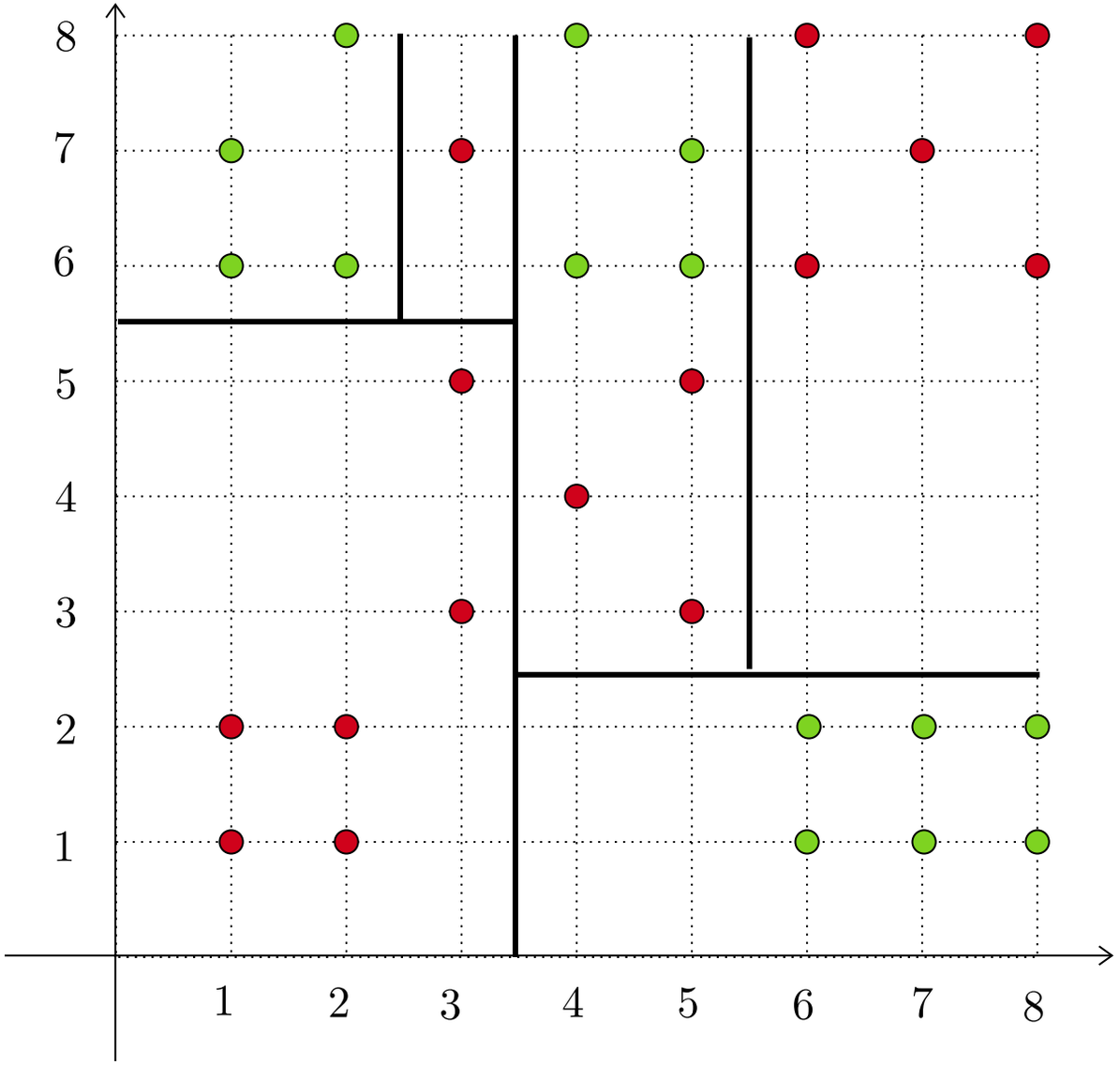


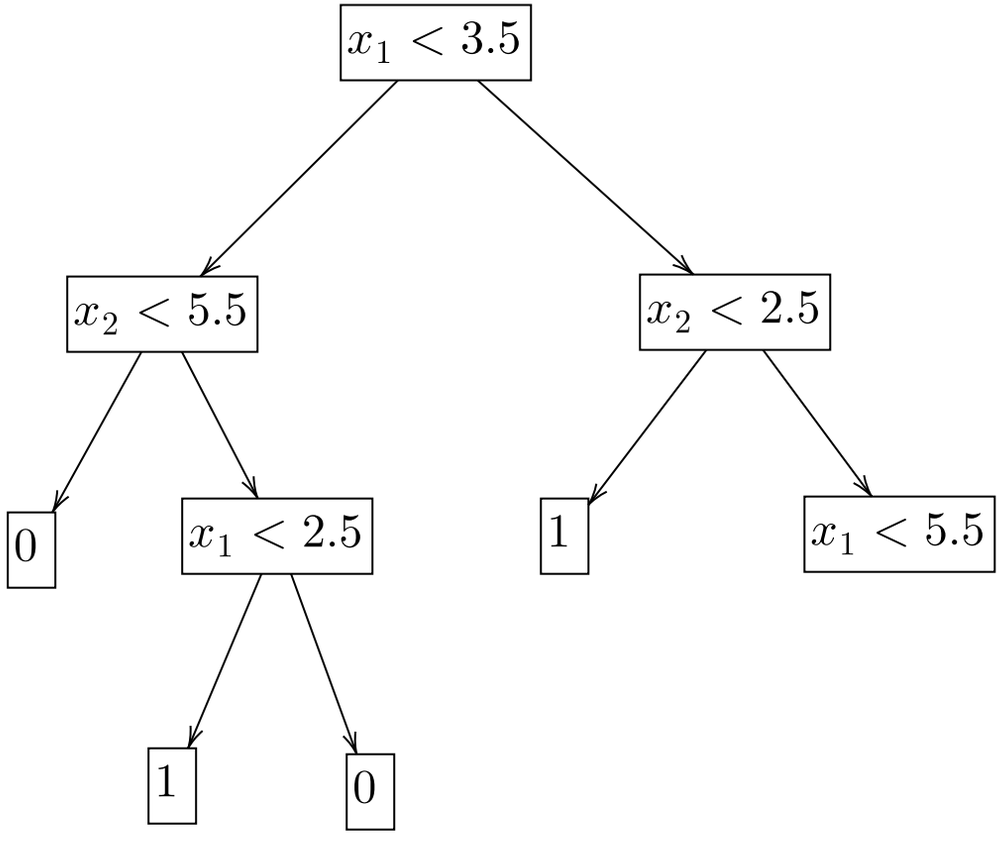
Growing a Tree



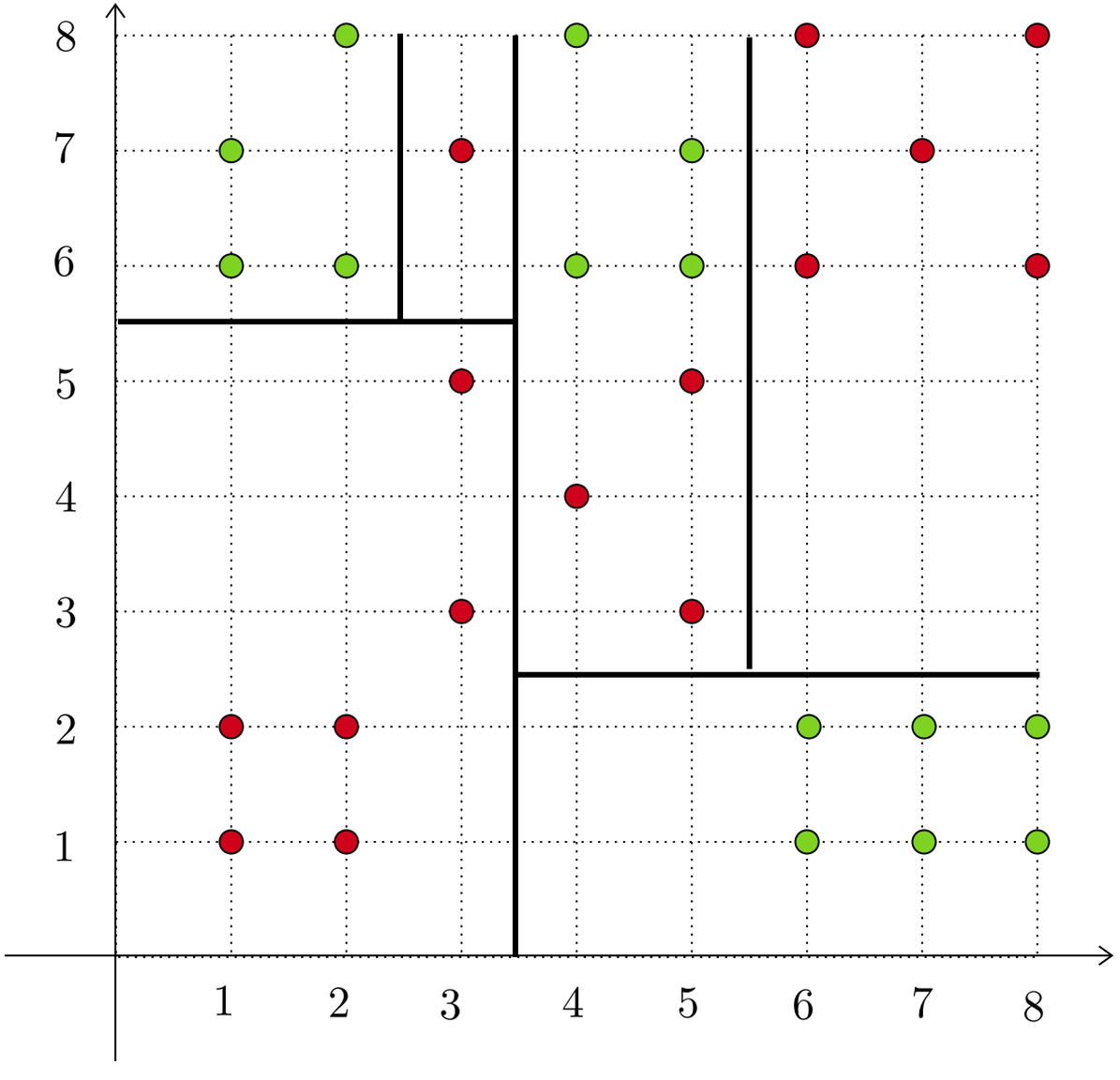


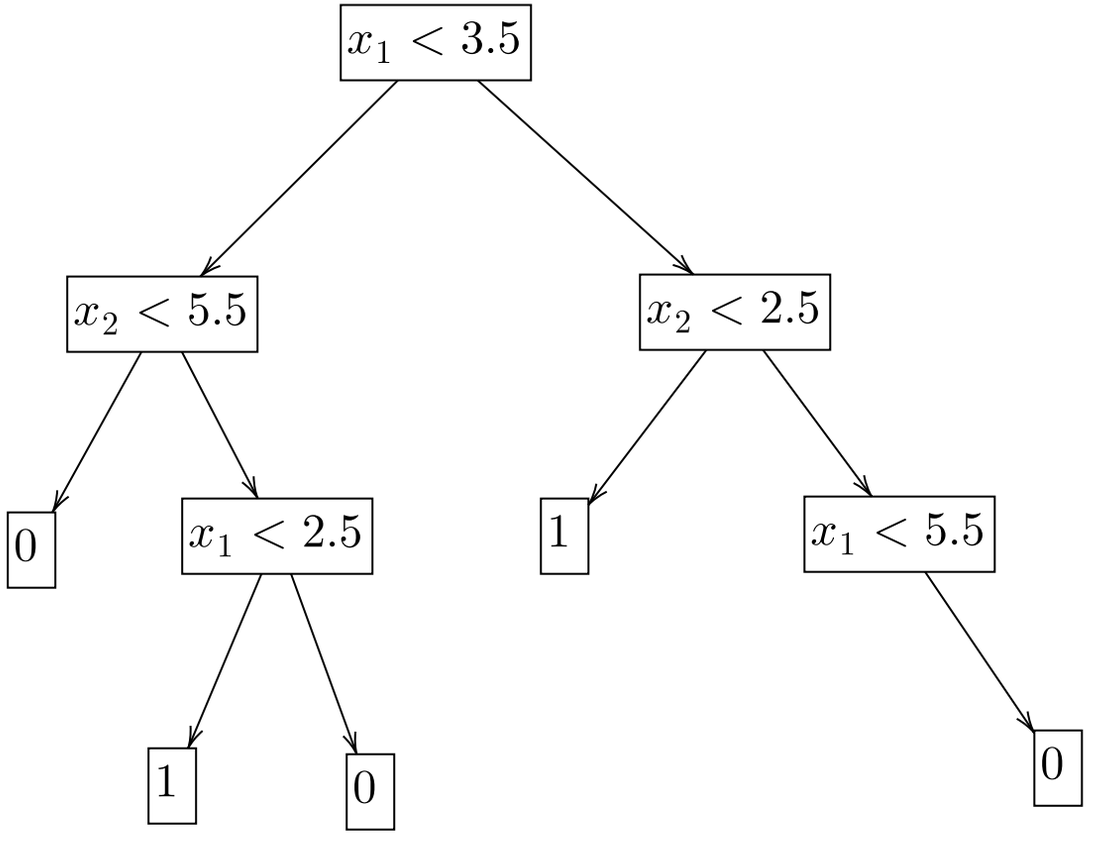
Growing a Tree



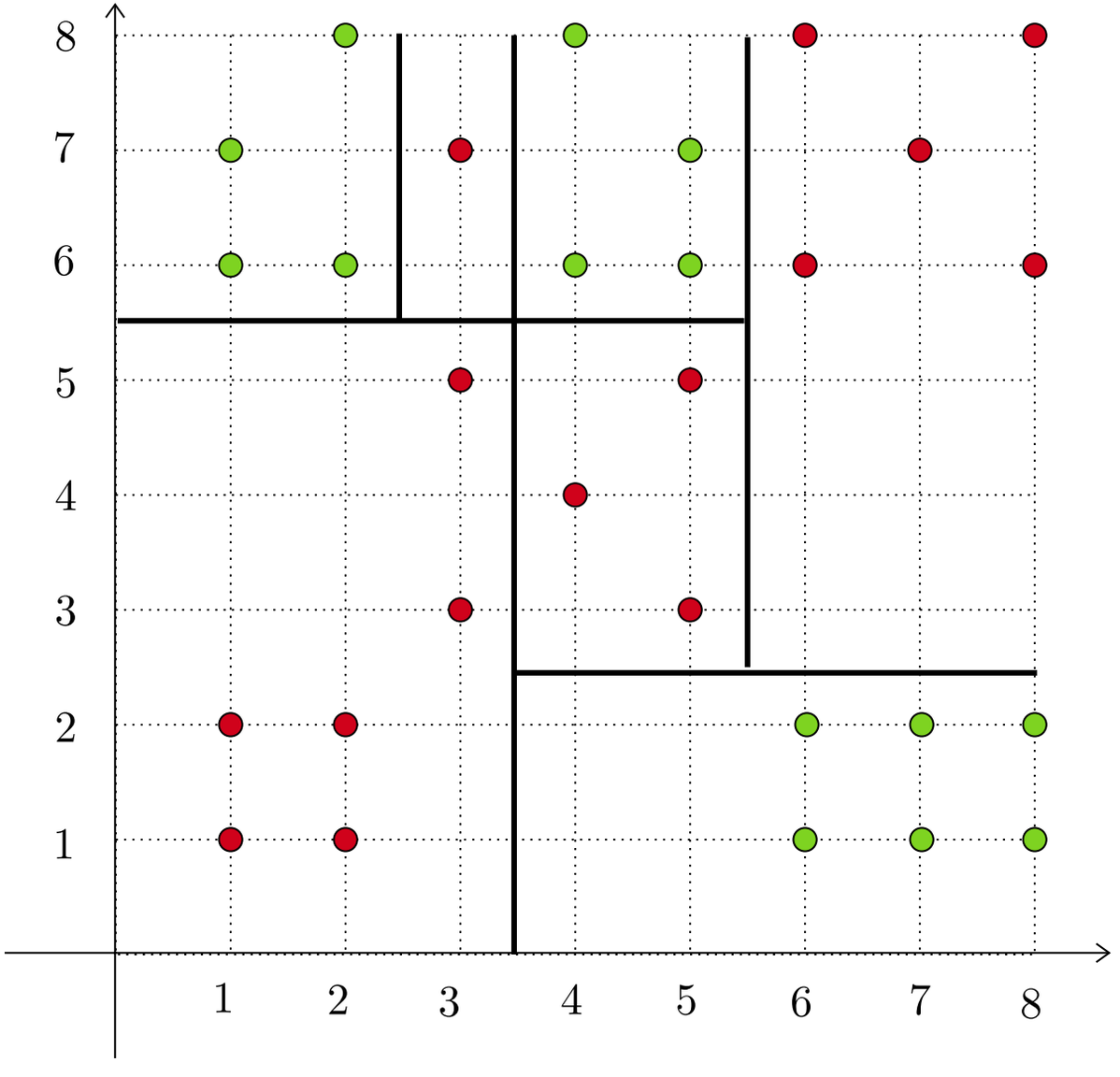


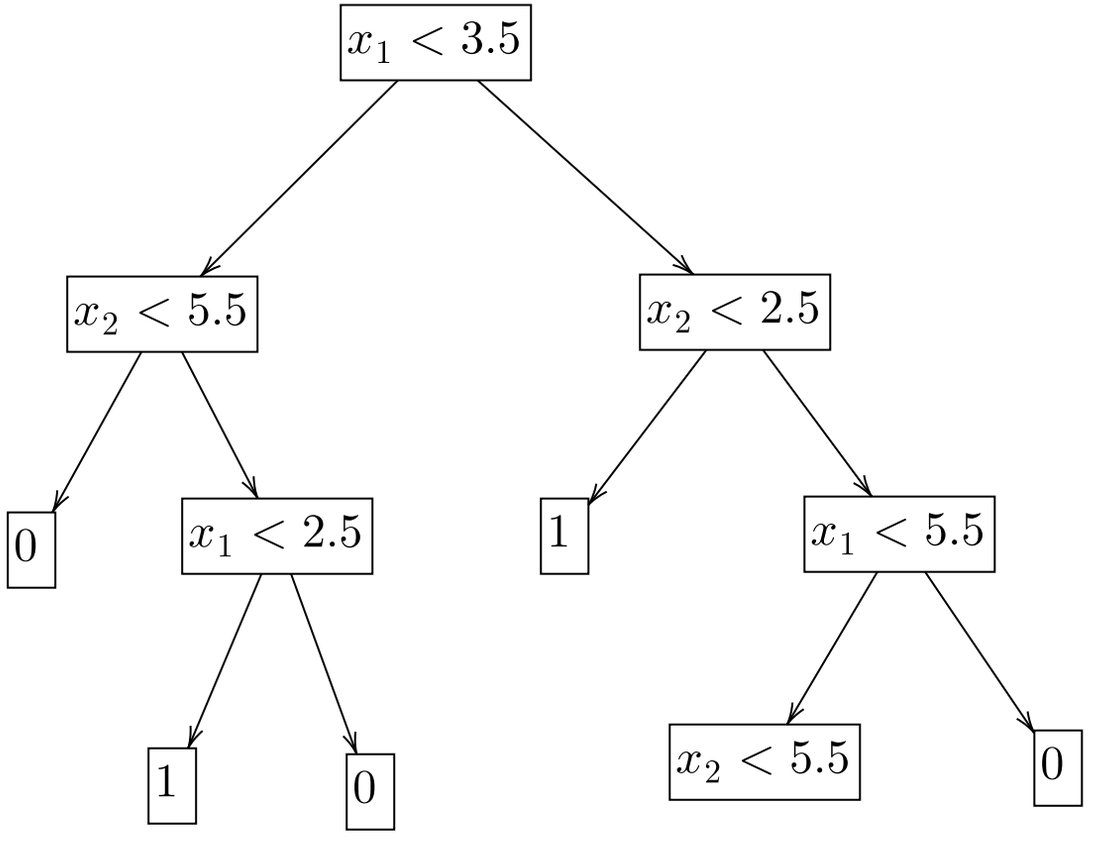
Growing a Tree



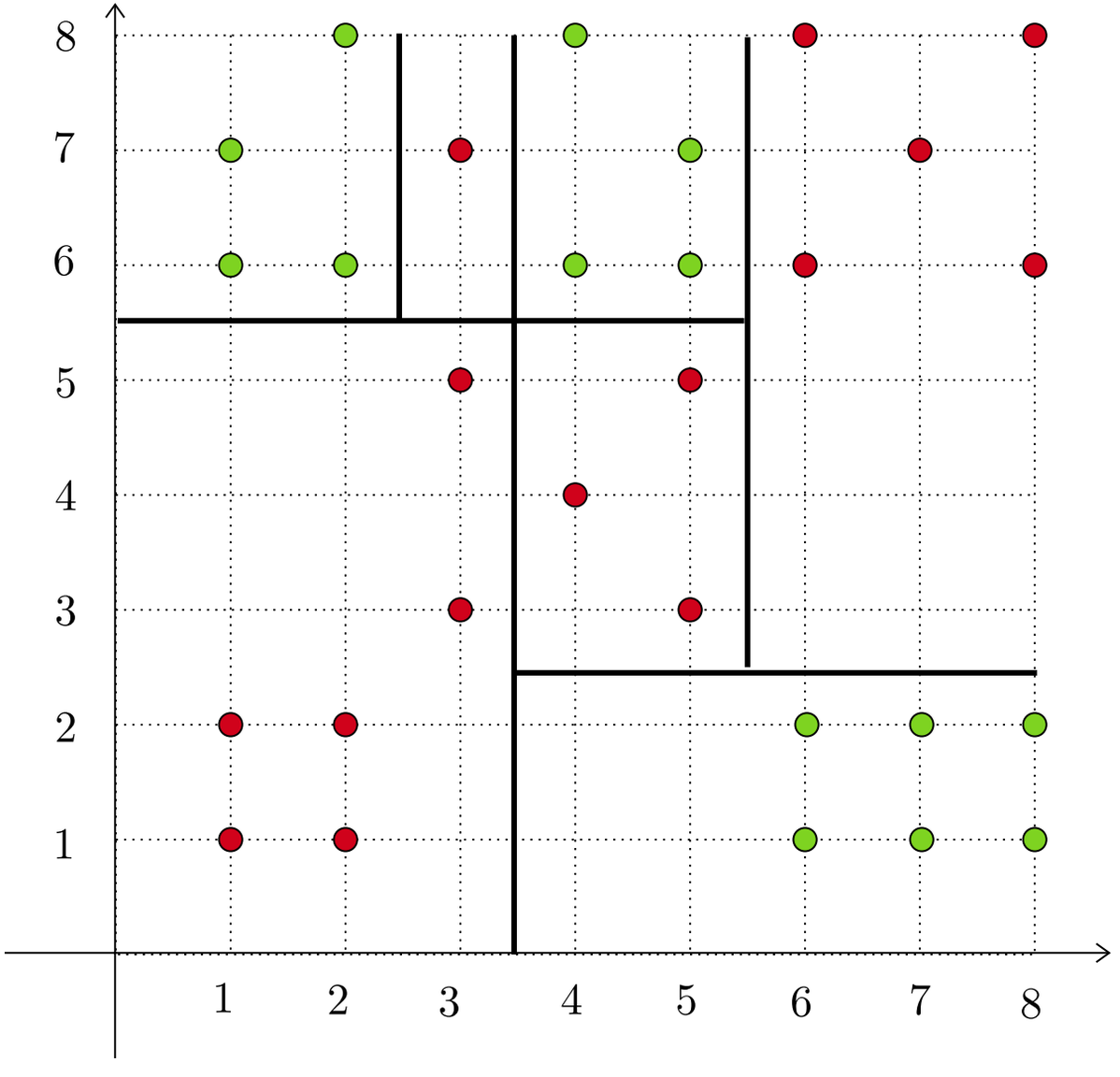


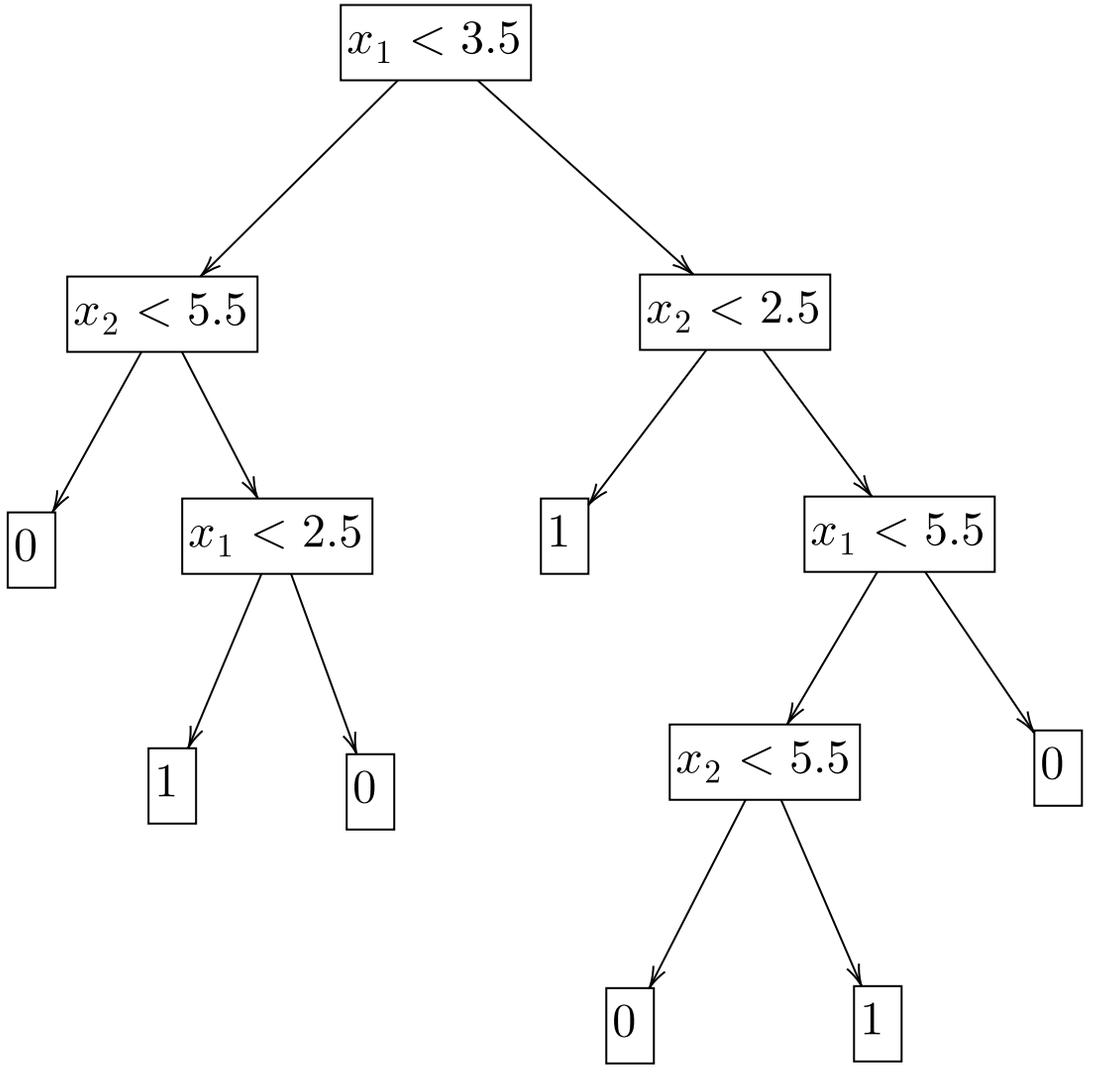
Growing a Tree



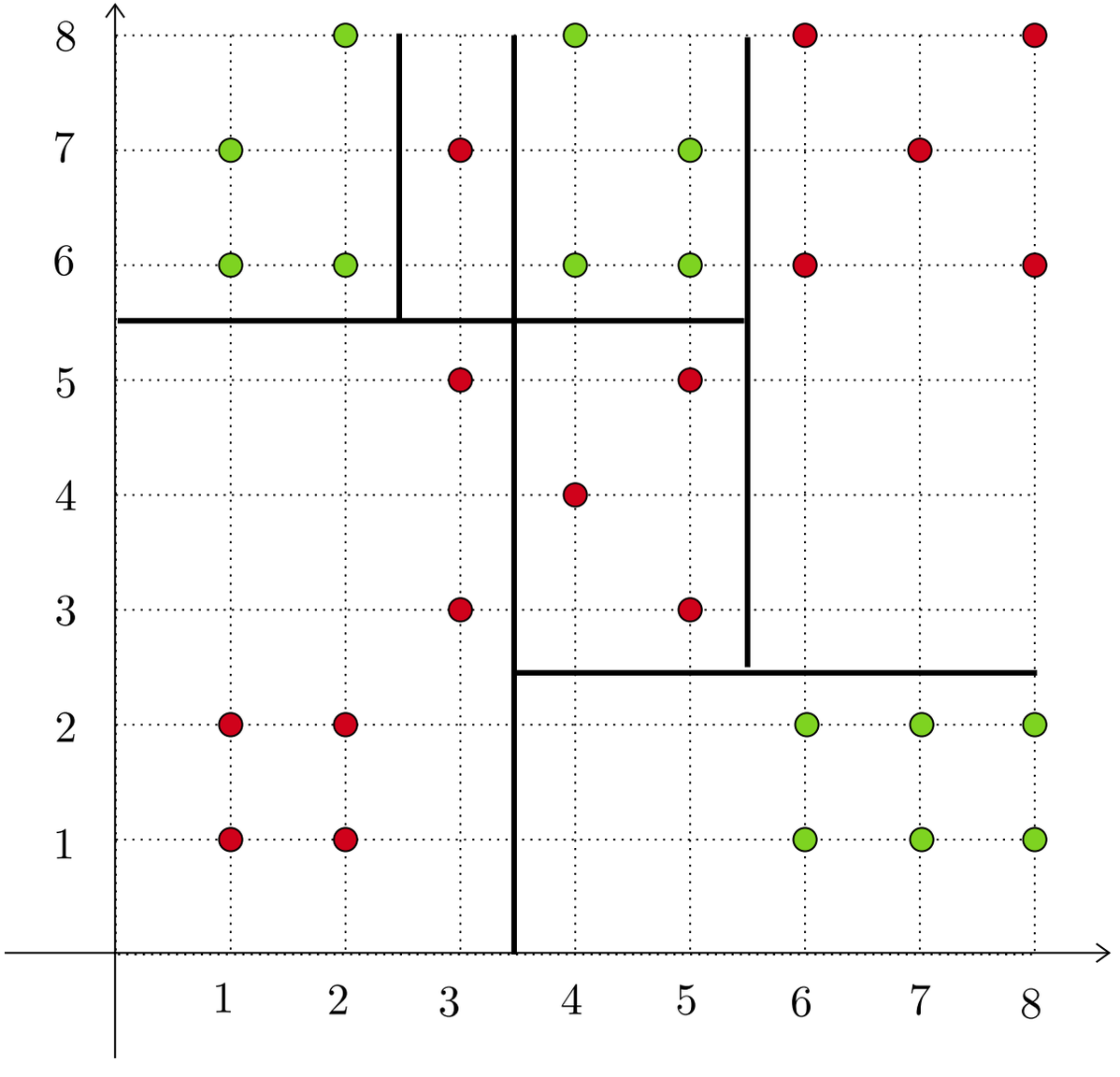


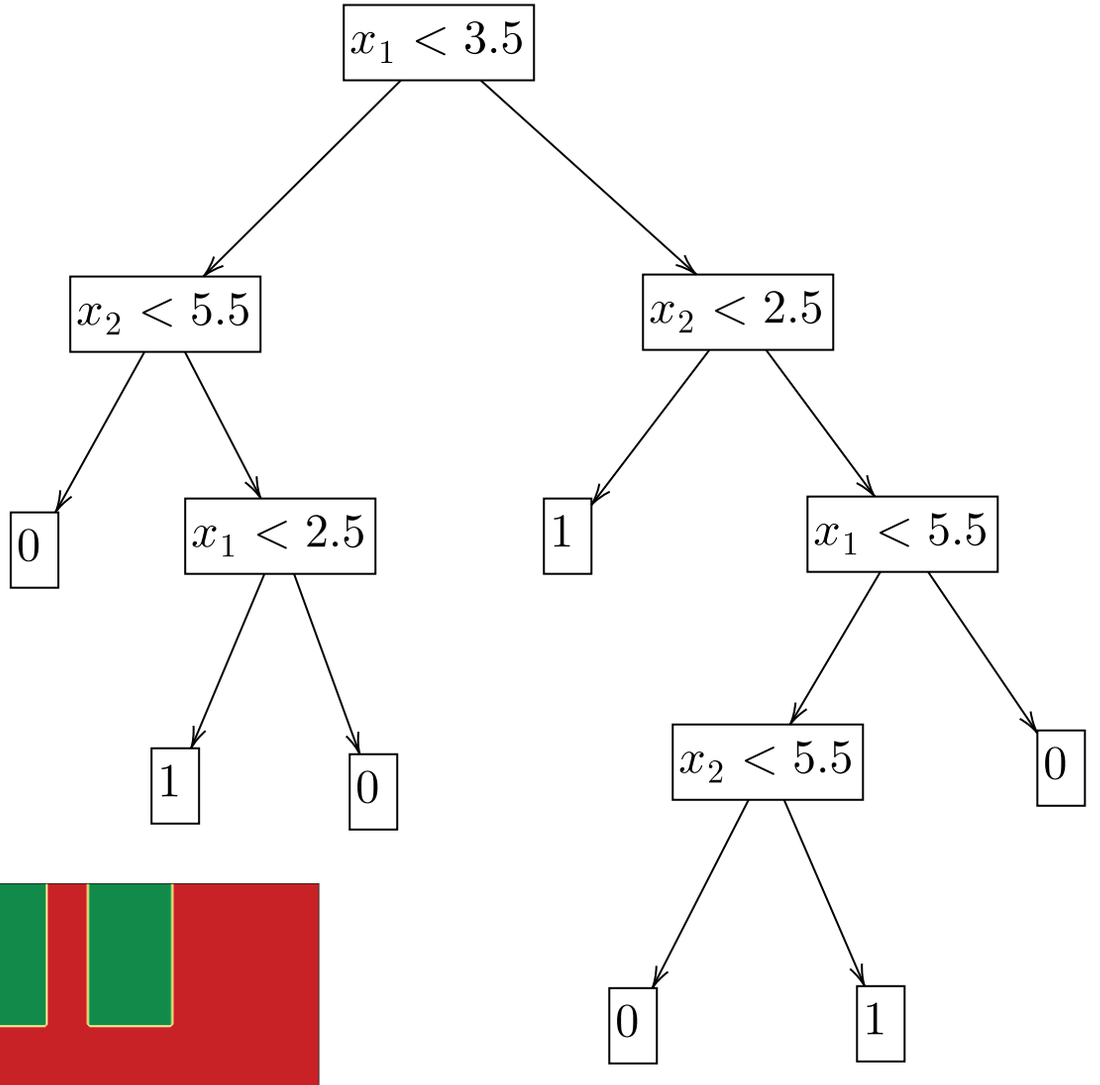
Growing a Tree

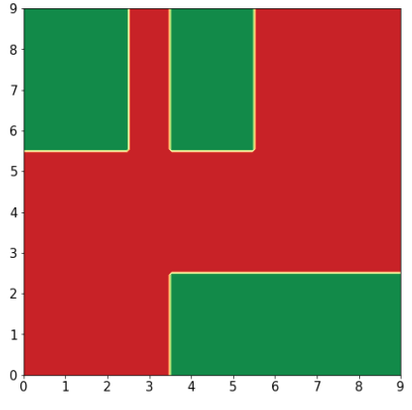




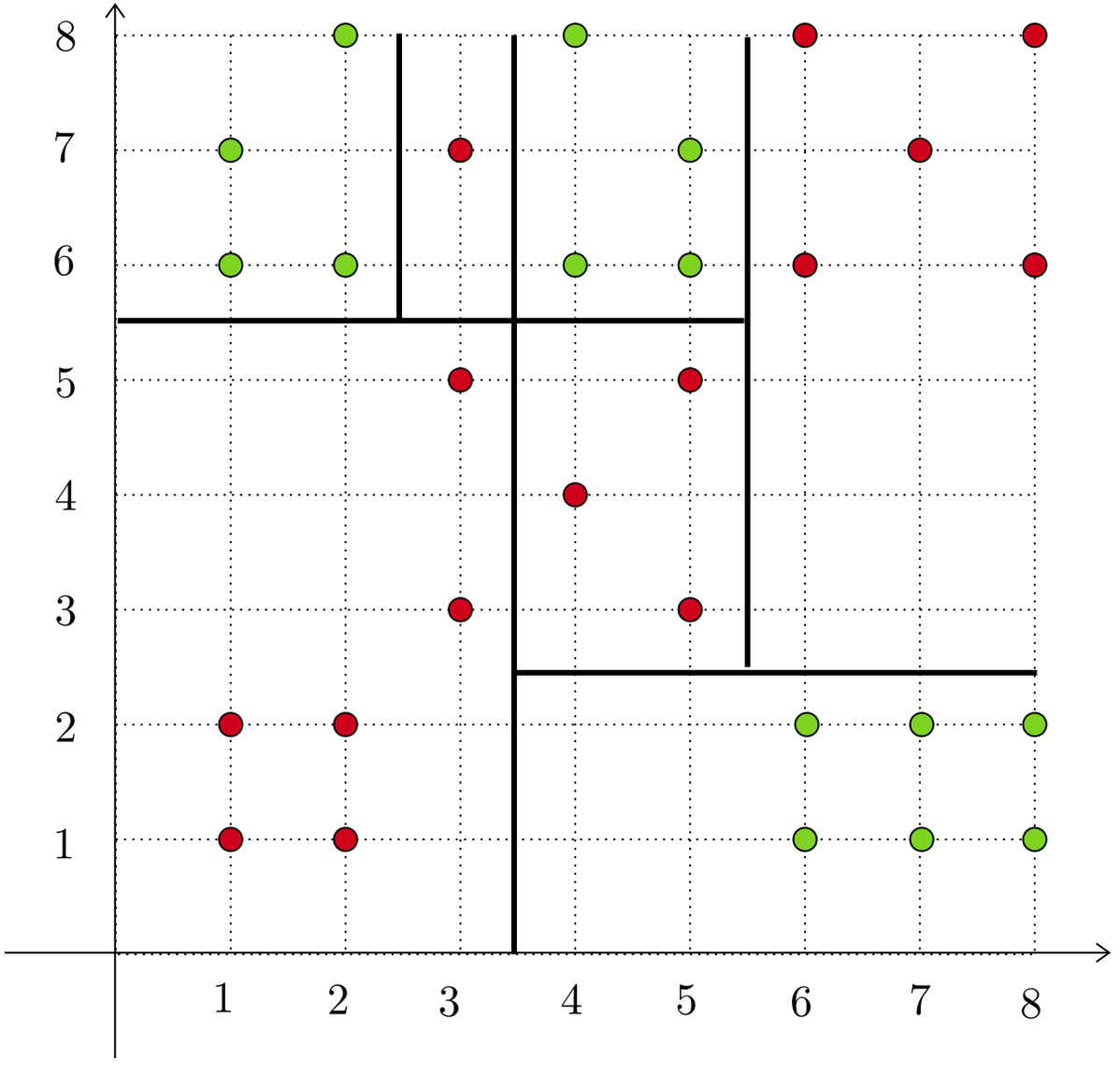
Growing a Tree

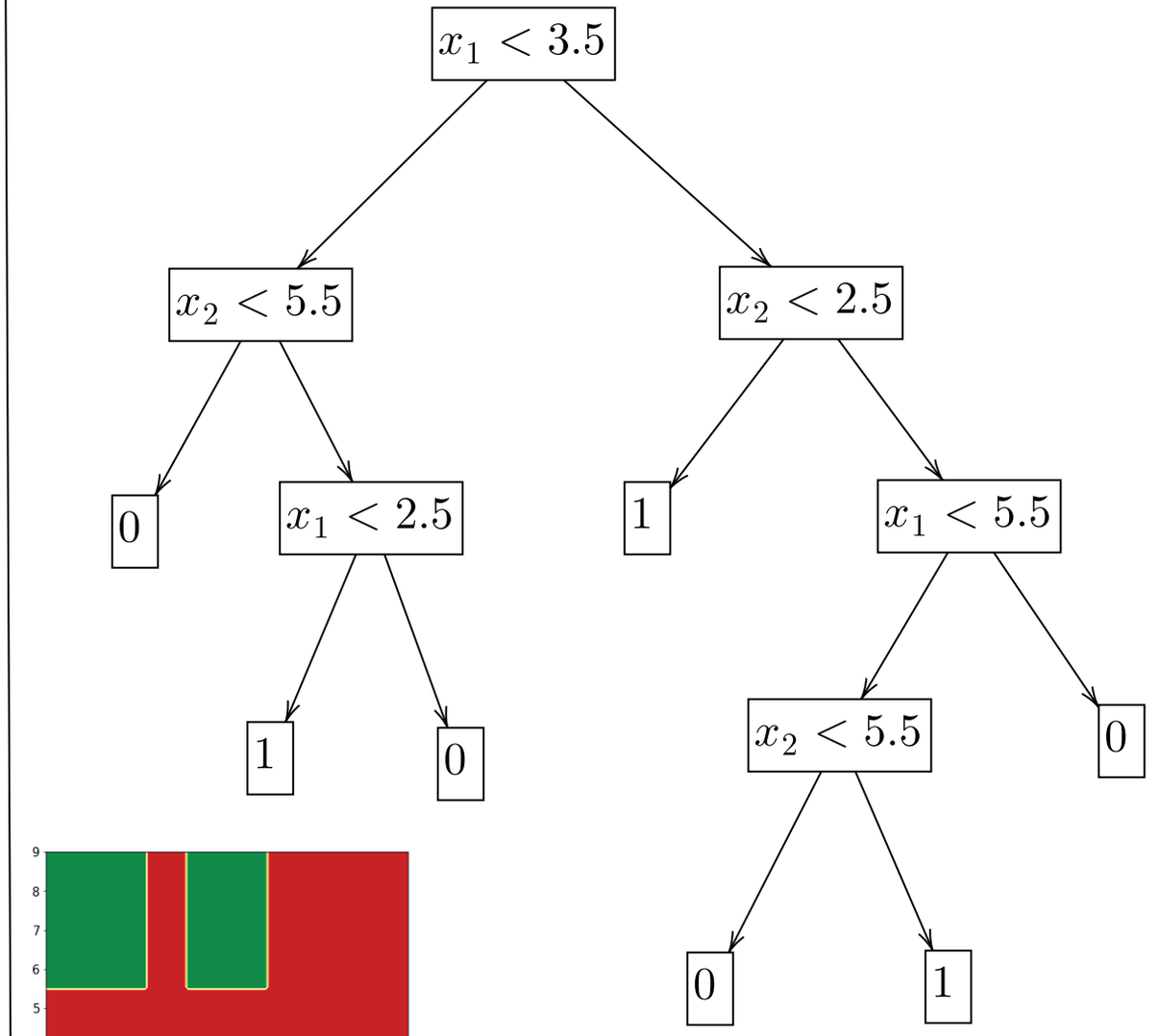


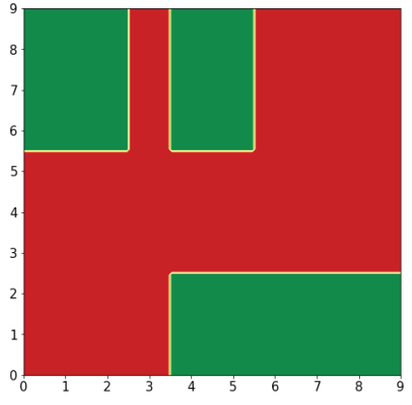




Growing a Tree







- 7 regions
- 7 leaves

Stopping Criteria and Pruning

Stopping Criterion

Stopping Criteria and Pruning

Stopping Criterion

- Leaves are pure (default)

Stopping Criteria and Pruning

Stopping Criterion

- Leaves are pure (default)

Pre-Pruning

- Minumum samples at leaf node

Stopping Criteria and Pruning

Stopping Criterion

- Leaves are pure (default)

Pre-Pruning

- Minumum samples at leaf node
- Maximum depth

Stopping Criteria and Pruning

Stopping Criterion

- Leaves are pure (default)

Pre-Pruning

- Minumum samples at leaf node
- Maximum depth
- Minimum decrease in impurity

Stopping Criteria and Pruning

Stopping Criterion

- Leaves are pure (default)

Pre-Pruning

- Minumum samples at leaf node
- Maximum depth
- Minimum decrease in impurity

Post-Pruning

Stopping Criteria and Pruning

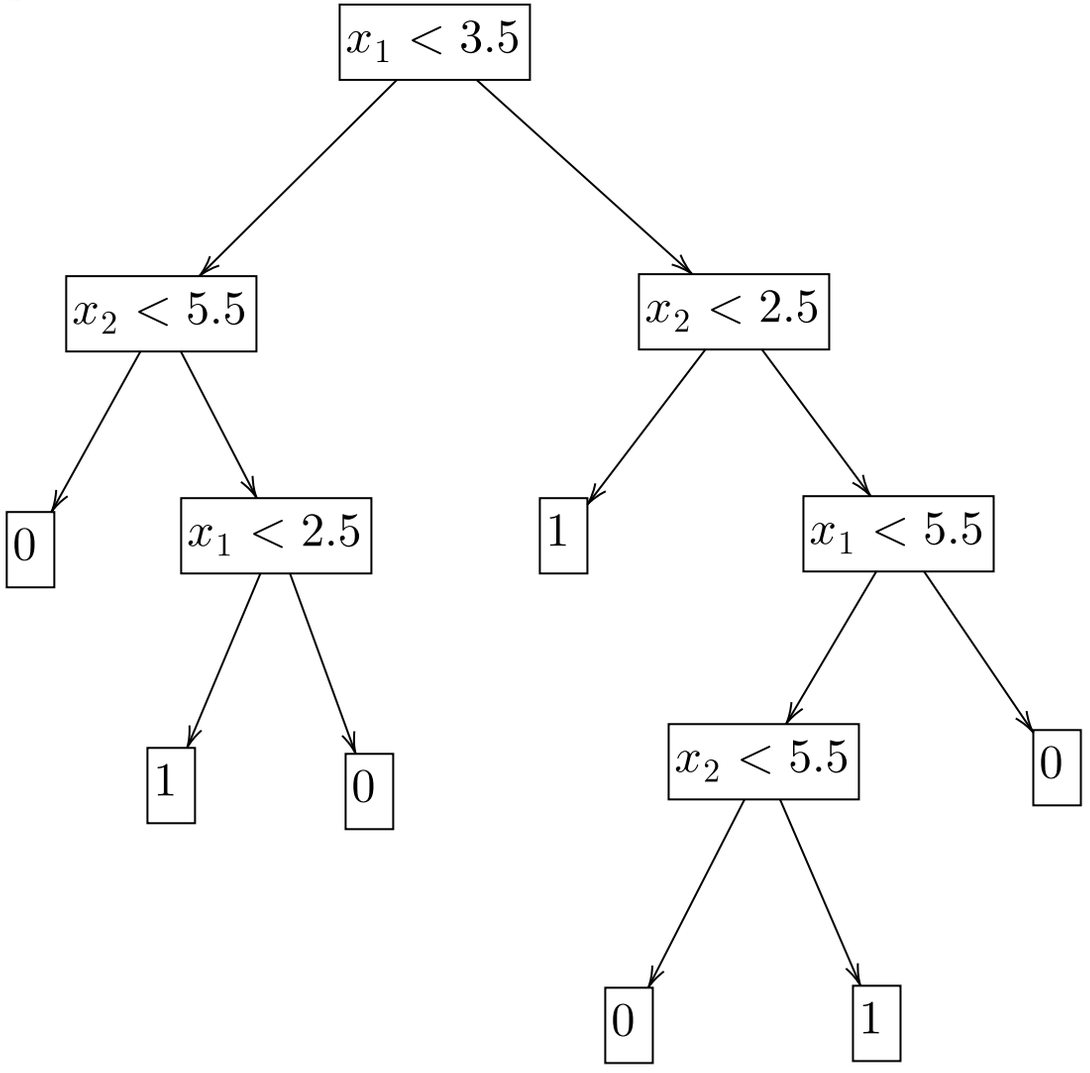
Stopping Criterion

- Leaves are pure (default)

Pre-Pruning

- Minumum samples at leaf node
- Maximum depth
- Minimum decrease in impurity

Post-Pruning



Stopping Criteria and Pruning

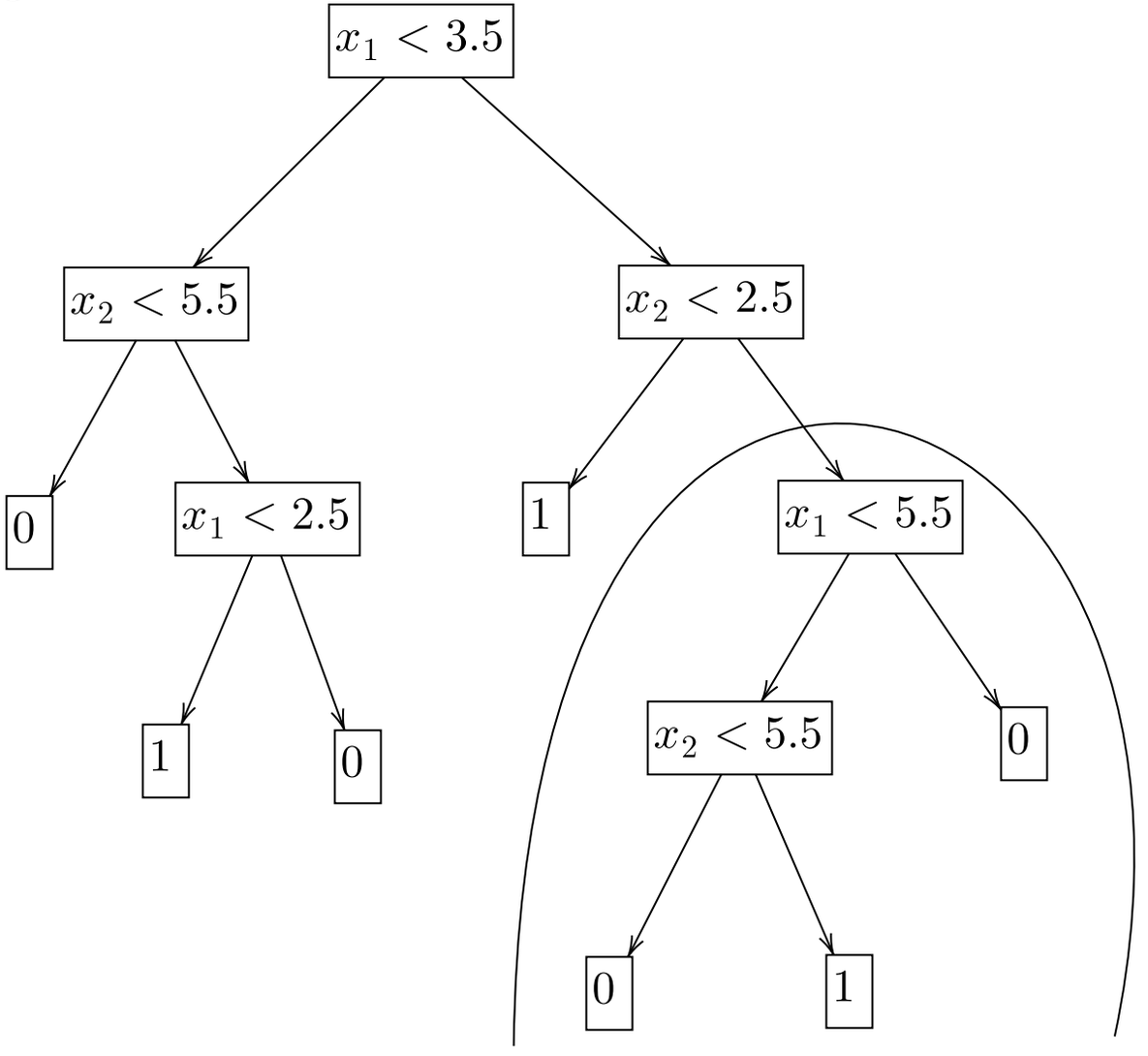
Stopping Criterion

- Leaves are pure (default)

Pre-Pruning

- Minumum samples at leaf node
- Maximum depth
- Minimum decrease in impurity

Post-Pruning



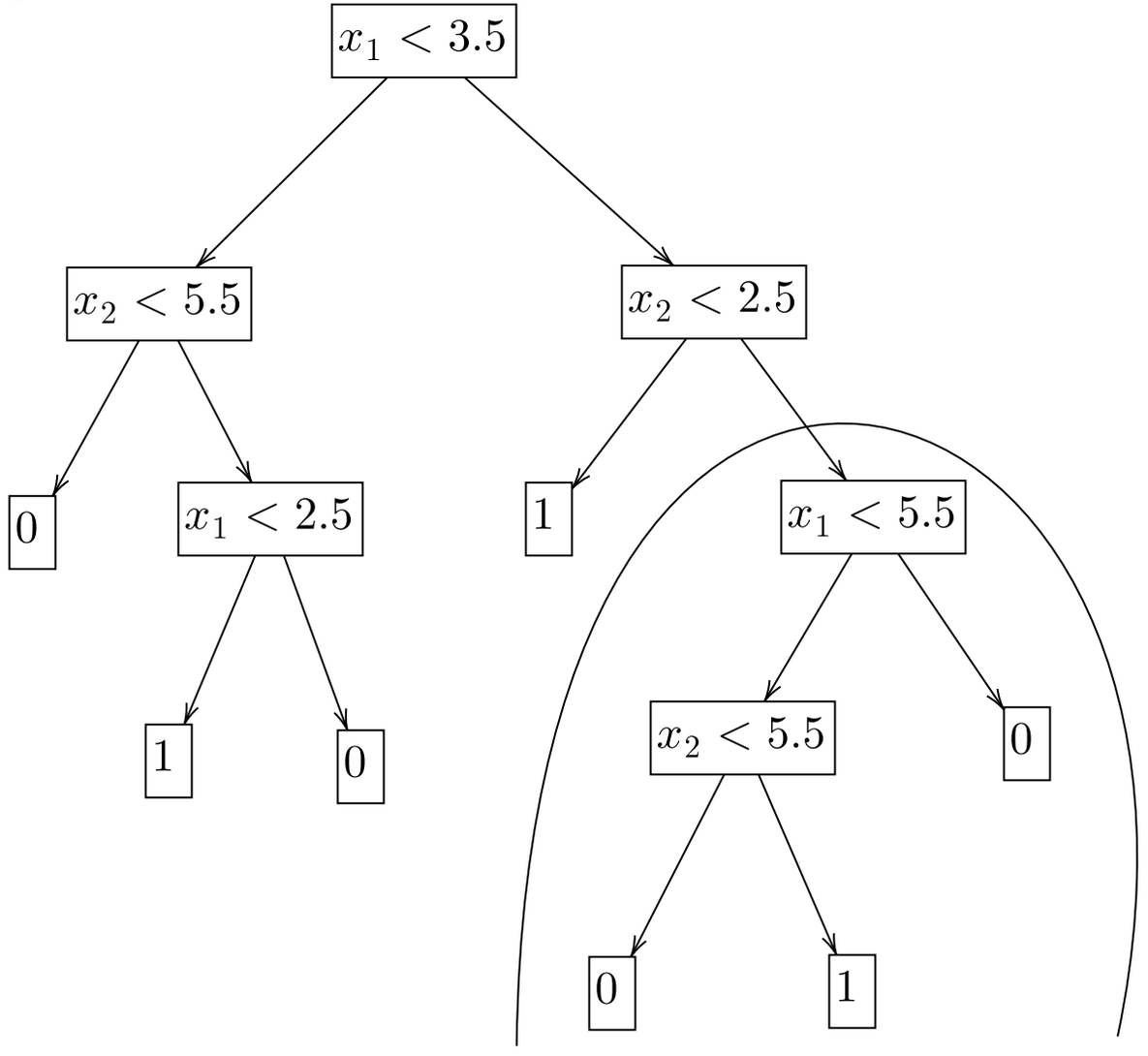
Stopping Criteria and Pruning

Stopping Criterion
- Leaves are pure (default)

Pre-Pruning
- Minumum samples at leaf node
- Maximum depth
- Minimum decrease in impurity

Post-Pruning
- Cost Complexity Pruning

$$\text{Cost} = \big(0 - 1 \text{ Loss}\big) + \lambda \cdot \big(\text{Num of leaves}\big)$$

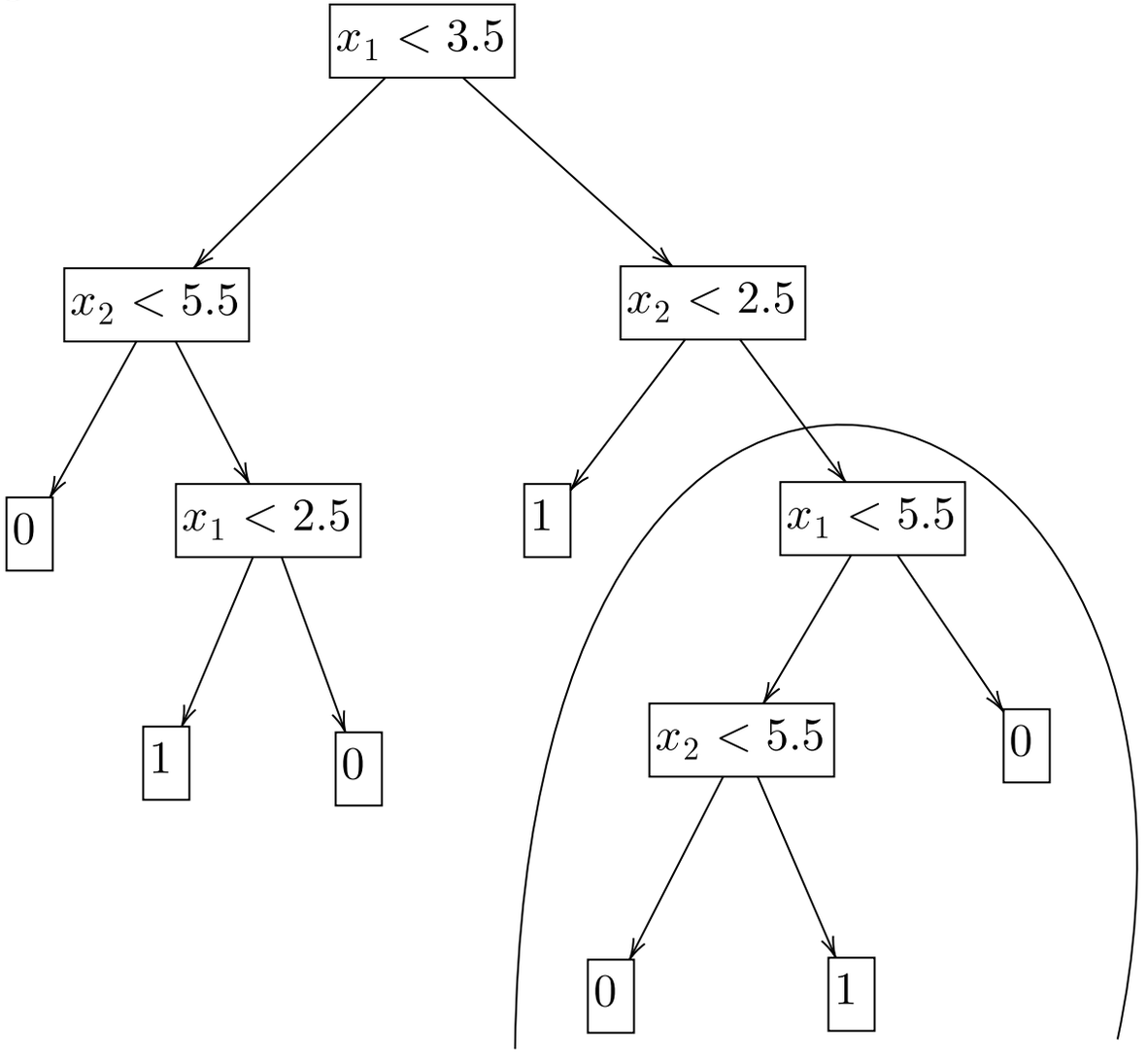


Stopping Criteria and Pruning

Stopping Criterion
- Leaves are pure (default)

Pre-Pruning
- Minumum samples at leaf node
- Maximum depth
- Minimum decrease in impurity

Post-Pruning
- Cost Complexity Pruning
  - subtrees



$$\text{Cost} = \big(0 - 1\ \text{Loss}\big) + \lambda \cdot \big(\text{Num of leaves}\big)$$

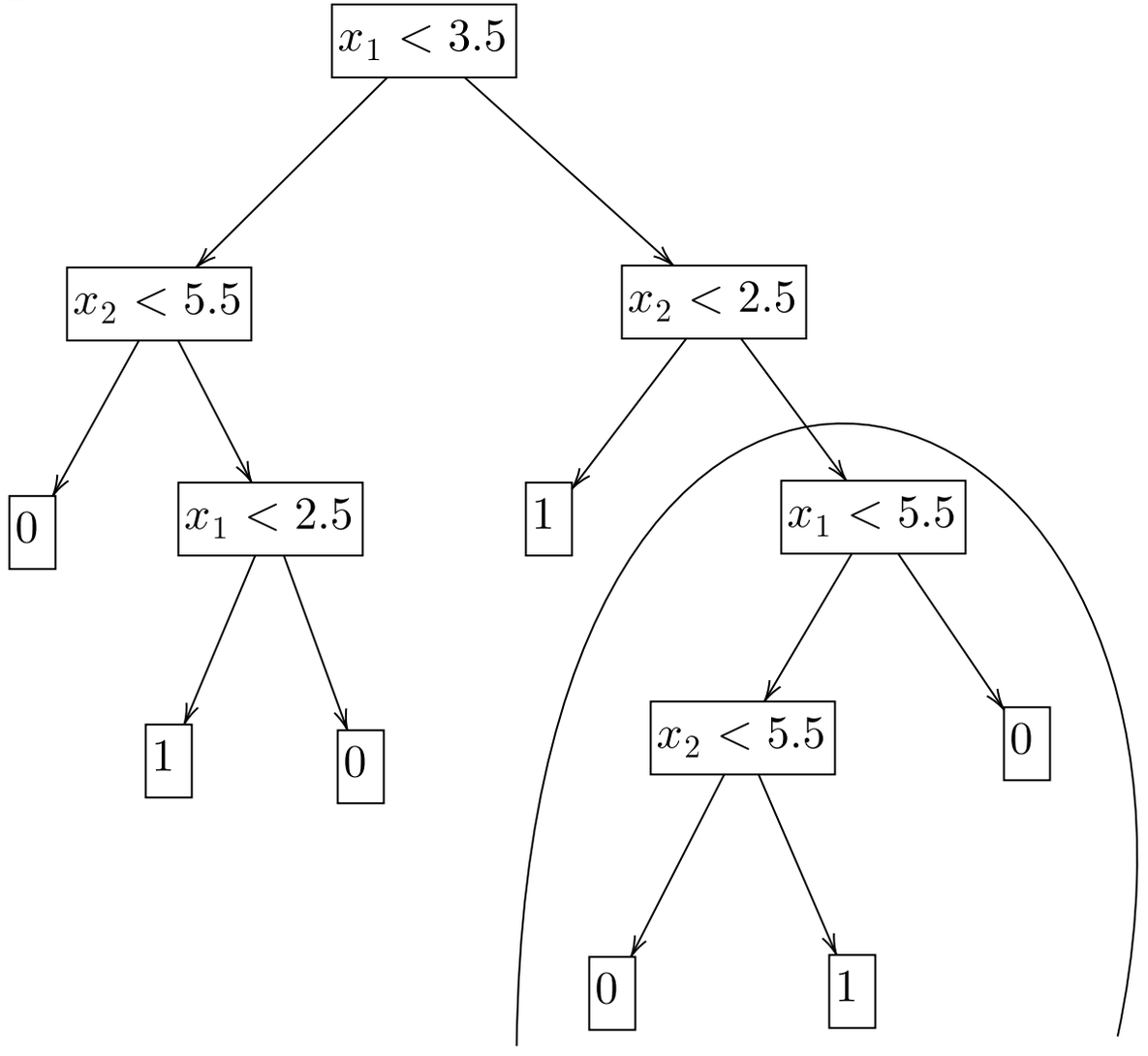
Stopping Criteria and Pruning

Stopping Criterion
- Leaves are pure (default)

Pre-Pruning
- Minumum samples at leaf node
- Maximum depth
- Minimum decrease in impurity

Post-Pruning
- Cost Complexity Pruning
  - subtrees
  - regularization

$$\text{Cost} = \big(0-1\ \text{Loss}\big) + \lambda \cdot \big(\text{Num of leaves}\big)$$



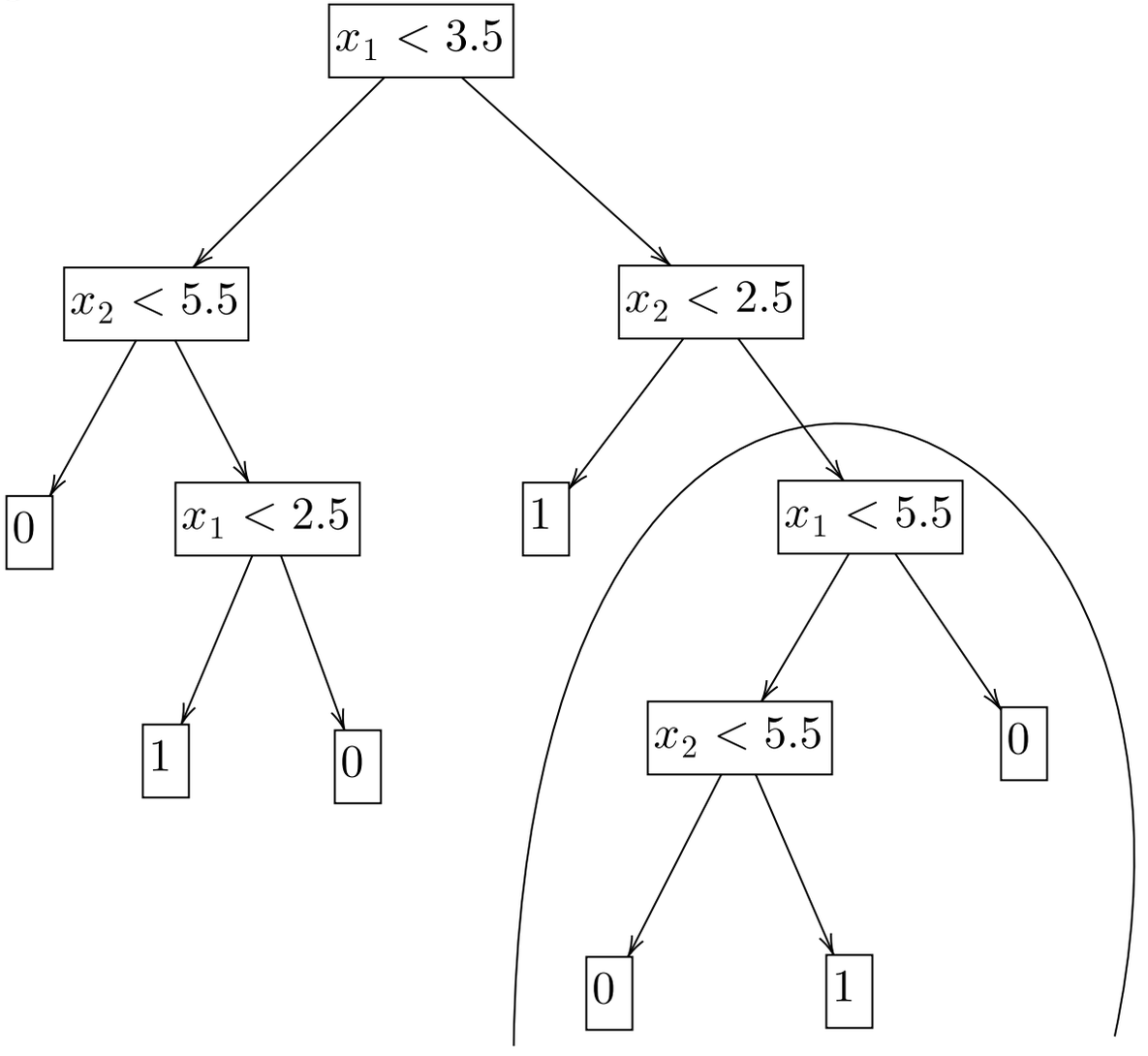
Stopping Criteria and Pruning

Stopping Criterion
- Leaves are pure (default)

Pre-Pruning
- Minumum samples at leaf node
- Maximum depth
- Minimum decrease in impurity

Post-Pruning
- Cost Complexity Pruning
  - subtrees
  - regularization
  - Use CV to estimate λ

$$\text{Cost} = \big(0-1\ \text{Loss}\big) + \lambda \cdot \big(\text{Num of leaves}\big)$$



Decision Regions

$\mathbb{R}$

Decision Regions

$\mathbb{R}$

Decision Regions

$\mathbb{R}$

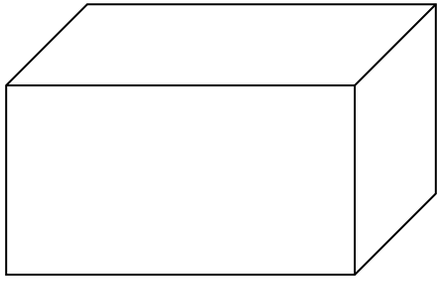
$\mathbb{R}^2$

Decision Regions

$\mathbb{R}$

$\mathbb{R}^2$

$\mathbb{R}^3$



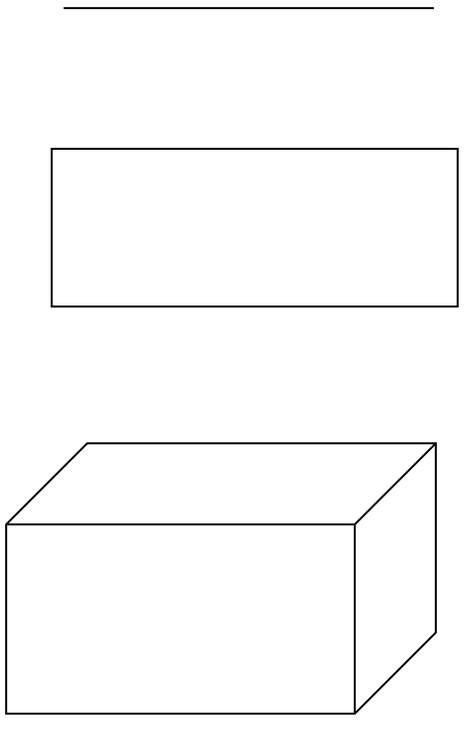
Decision Regions

$\mathbb{R}$

$\mathbb{R}^2$

$\mathbb{R}^3$

$\mathbb{R}^d$



hyper-rectangles

Advantanges

- Interpretable
- Can be displayed graphically
- Can be understood by non-experts

Disadvantages

Advantanges

- Interpretable
- Can be displayed graphically
- Can be understood by non-experts

Disadvantages

- Low predictive power
- High variance, i.e., sensitive to small changes in the training dataset

Misc Details

- Model:
- Search space:
- Type of algorithm:
- Used for:
  -
  -
- Terms:
  -
  -
  -

Misc Details

- Model: tree
- Search space: trees
- Type of algorithm: greedy, top-down
- Used for:
  - classification
  - regression
- Terms:
  - CART
  - ID3
  - C4.5