



Please Sit With Your Group

L-4

Supervised Learning III

MAIB7002

Fundamentals of AI Models and Technologies

Coding Support Sessions

Optional

Cyberport · 12:30–14:00

Date	Room
30 September	CP-EFG
7 October	CP-EFG
14 October	CP-ABC

Course code walkthroughs and Q&A

Self-Study Resources

Focus	Resource	Suggested Topics
Python Basics	Michigan: Programming for Everybody https://www.coursera.org/learn/python	Variables, Functions and Loops
NumPy and Matplotlib	Stanford: Python NumPy Tutorial https://cs231n.github.io/python-numpy-tutorial/	Arrays, Indexing and Plotting
Linear Algebra	3Blue1Brown: Visual Introduction https://www.3blue1brown.com/topics/linear-algebra	Vectors, Matrices and Dot Products
Course Reading	Raschka, Liu and Mirjalili (2022) <i>Machine Learning with PyTorch and Scikit-Learn</i>	Ch 1–3: Foundations and Supervised Learning Ch 6: Model Evaluation

Course Notebook Updates

maib7002.dirk.hk/L1_v2.zip · maib7002.dirk.hk/L2_v2.zip · maib7002.dirk.hk/L3_v2.zip

Label Inside Code Cells

What to Do

CORE: UNDERSTAND

Explain the inputs, operation and output

CORE: MODIFY

Make the requested edit, rerun and explain

REFERENCE: RUN AS SUPPLIED

Run the supplied code; interpret the indicated outputs

Open [start_here.ipynb](#) for guidance · Same computations and data

Suggestion: Use an Agent With Tools

Chatbot

ChatGPT · Claude · Gemini · Qwen

Agent With Tools

Codex · Claude Code · Gemini CLI · Qwen Code

Suggestion: Use an Agent With Tools

Chatbot

ChatGPT · Claude · Gemini · Qwen

Agent With Tools

Codex · Claude Code · Gemini CLI · Qwen Code

Demo: Get the Course Notebooks Running

How do I set it up?

Here are the steps



You carry out the steps

Suggestion: Use an Agent With Tools

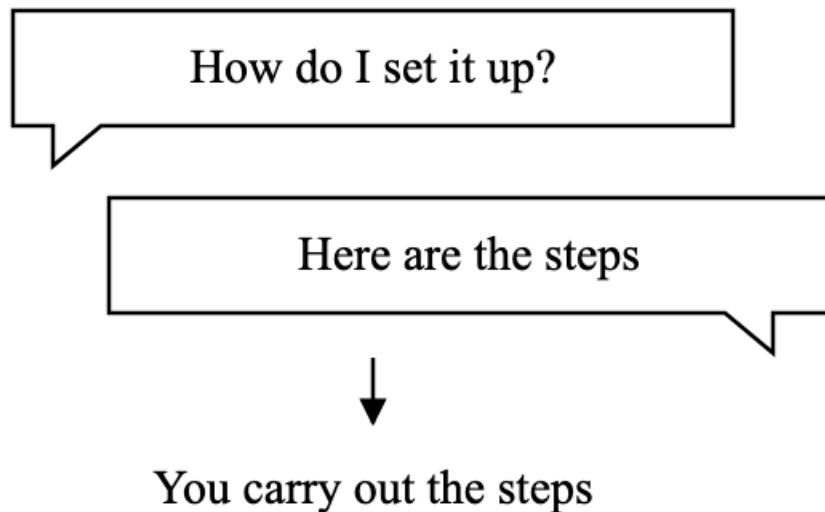
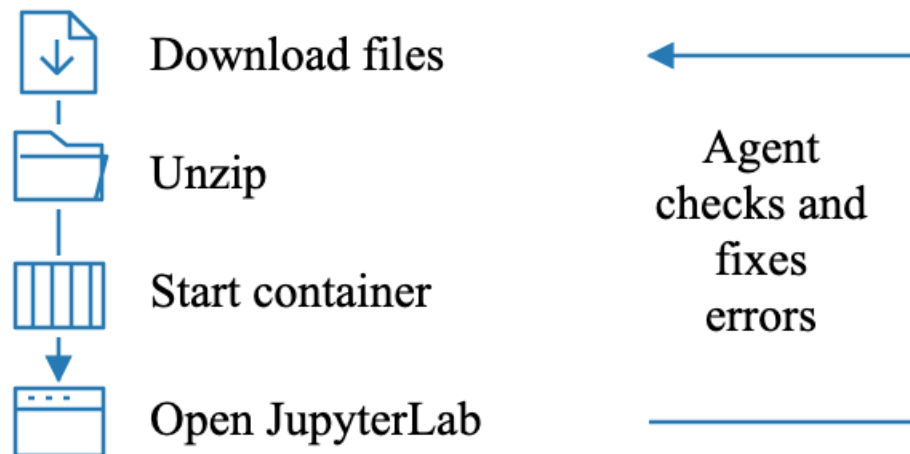
Chatbot

ChatGPT · Claude · Gemini · Qwen

Agent With Tools

Codex · Claude Code · Gemini CLI · Qwen Code

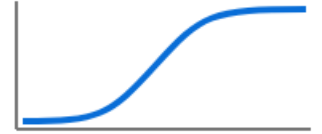
Demo: Get the Course Notebooks Running



Supervised Learning: L3 and L4

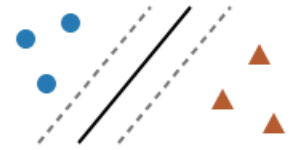
Logistic Regression

Estimate Class Probabilities



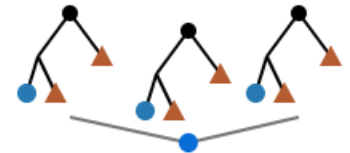
SVM

Separate Classes with a Margin



Random Forest

Learn Questions, Then Combine Trees

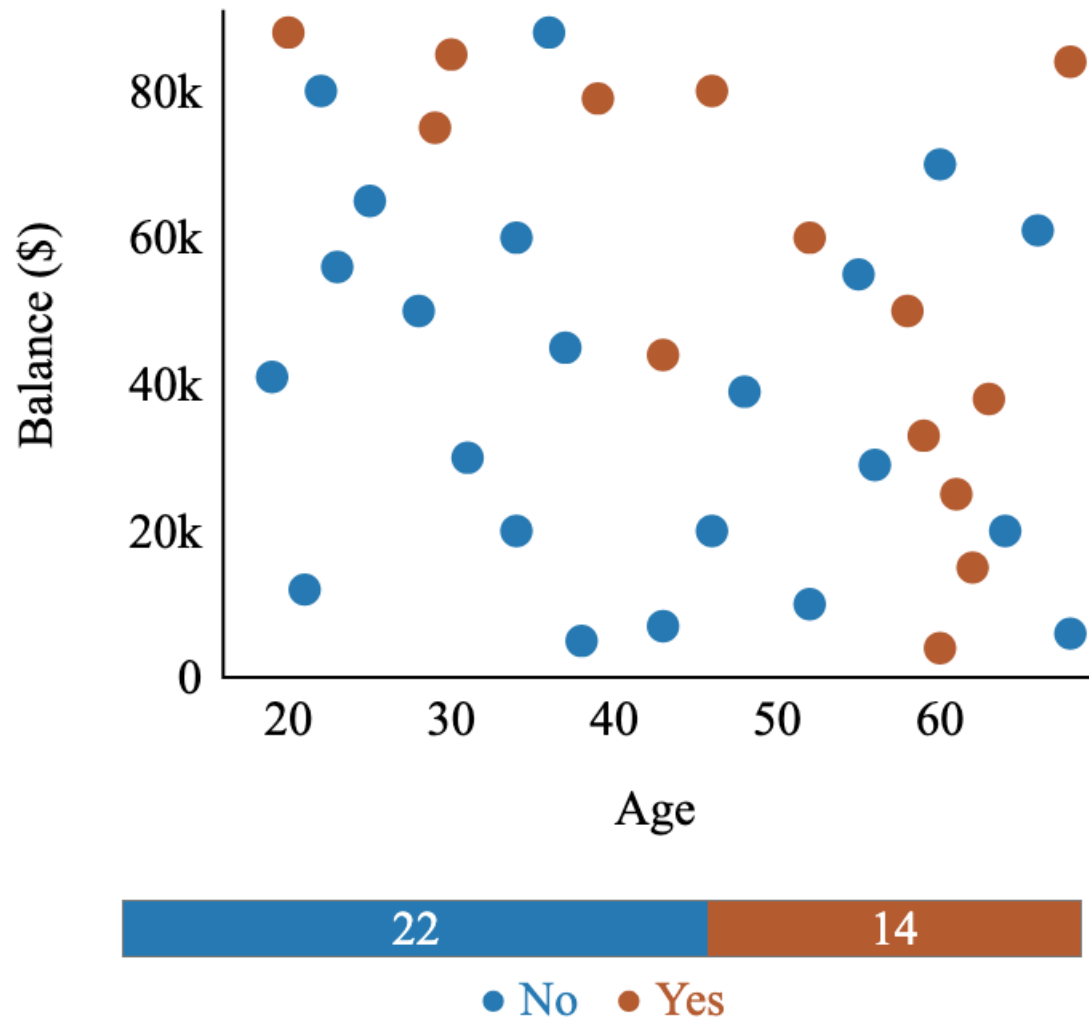


kNN

Use the Labels of Nearby Samples

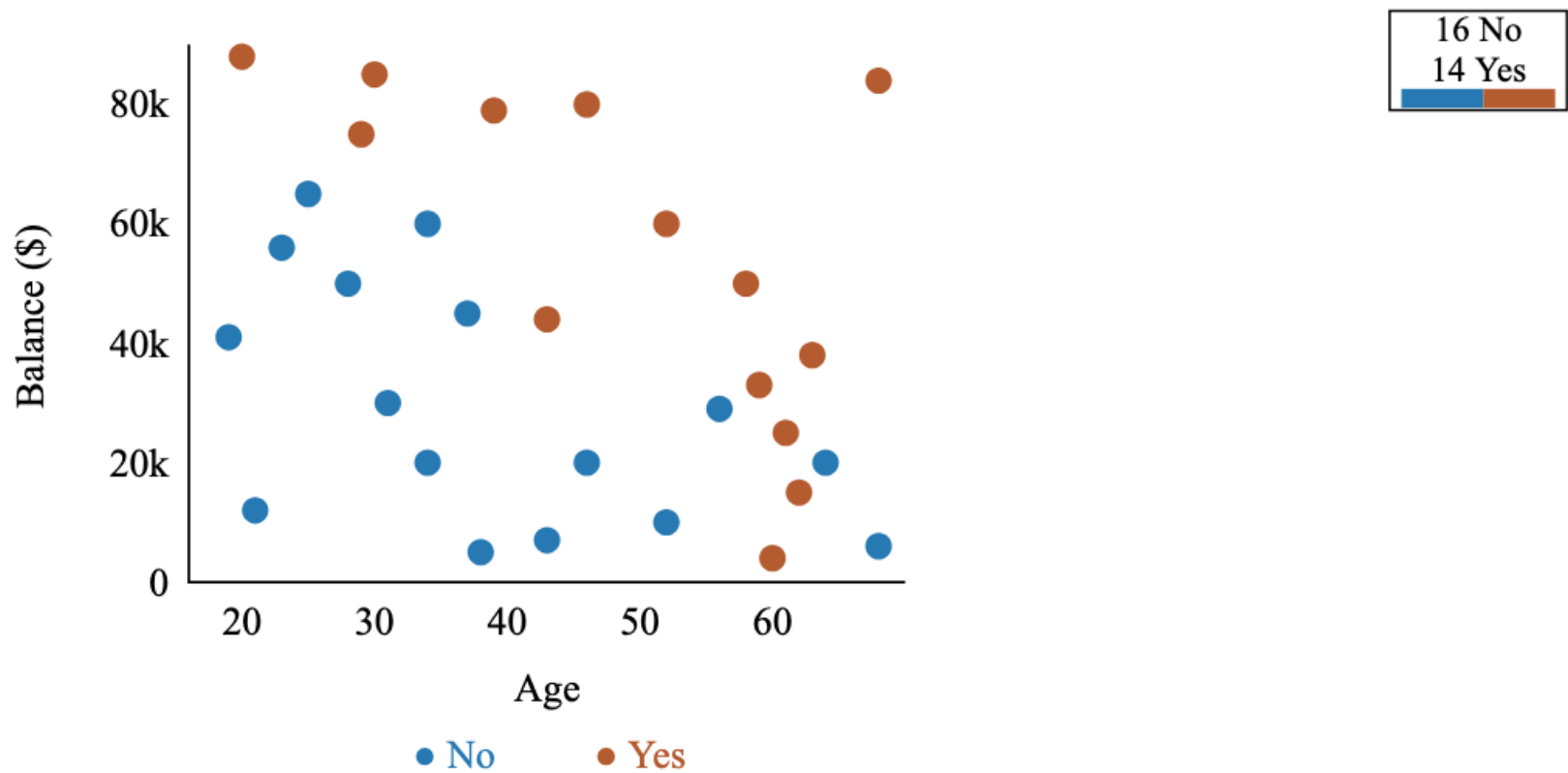


Who Will Open a Term Deposit?



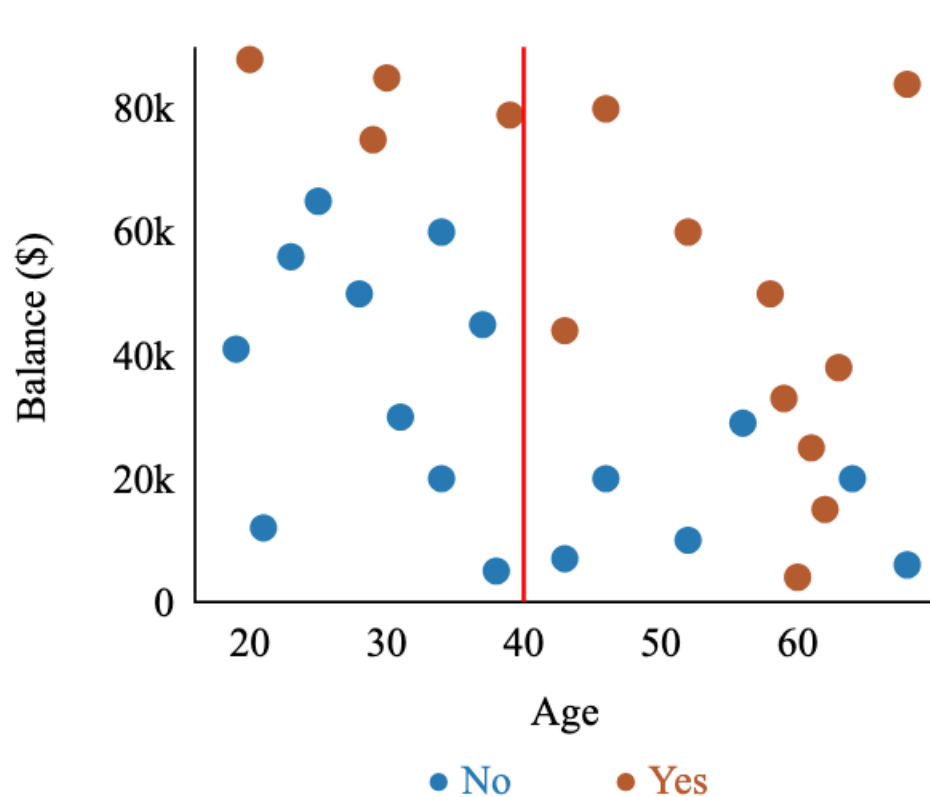
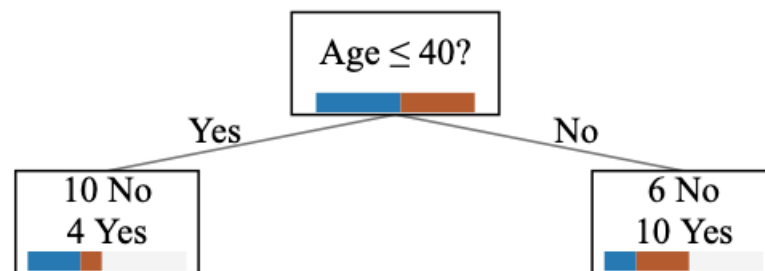
A Tree Divides Customers by Asking Questions

Example 1: A 30-Customer Subset



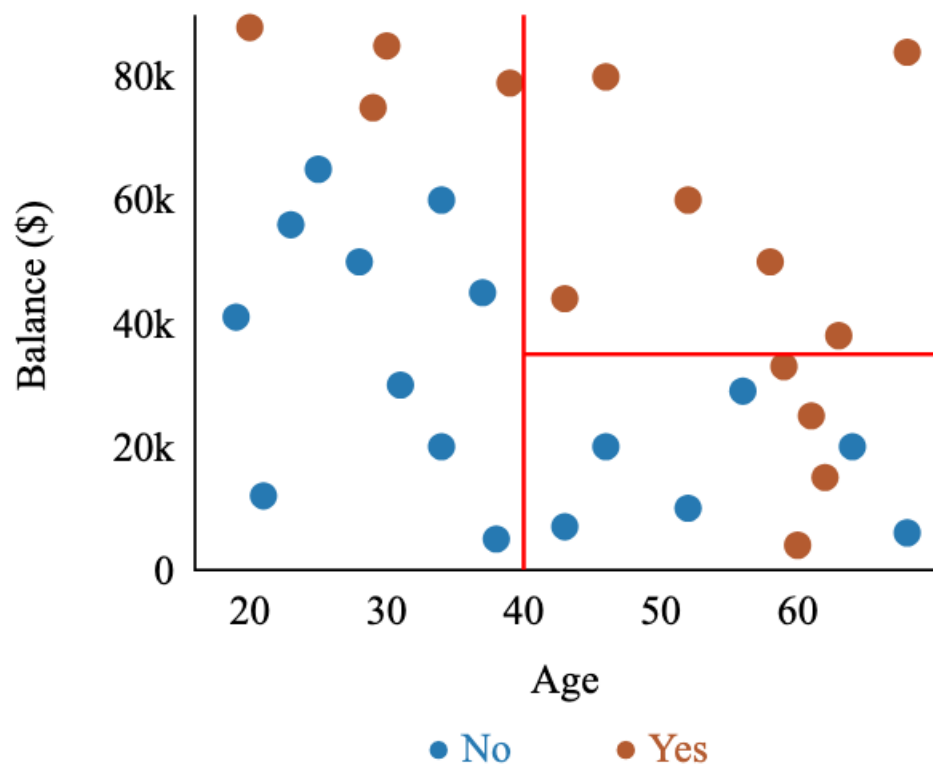
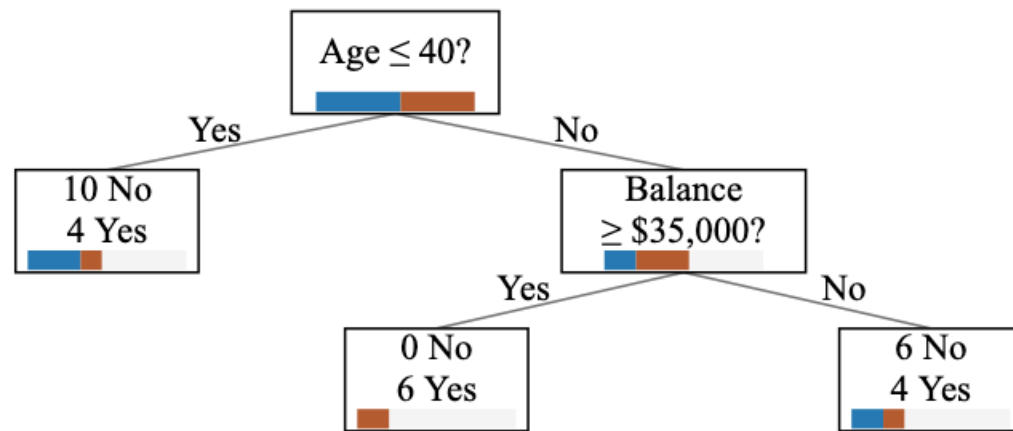
A Tree Divides Customers by Asking Questions

Ask About Age: Divide the Customers Into Two Groups



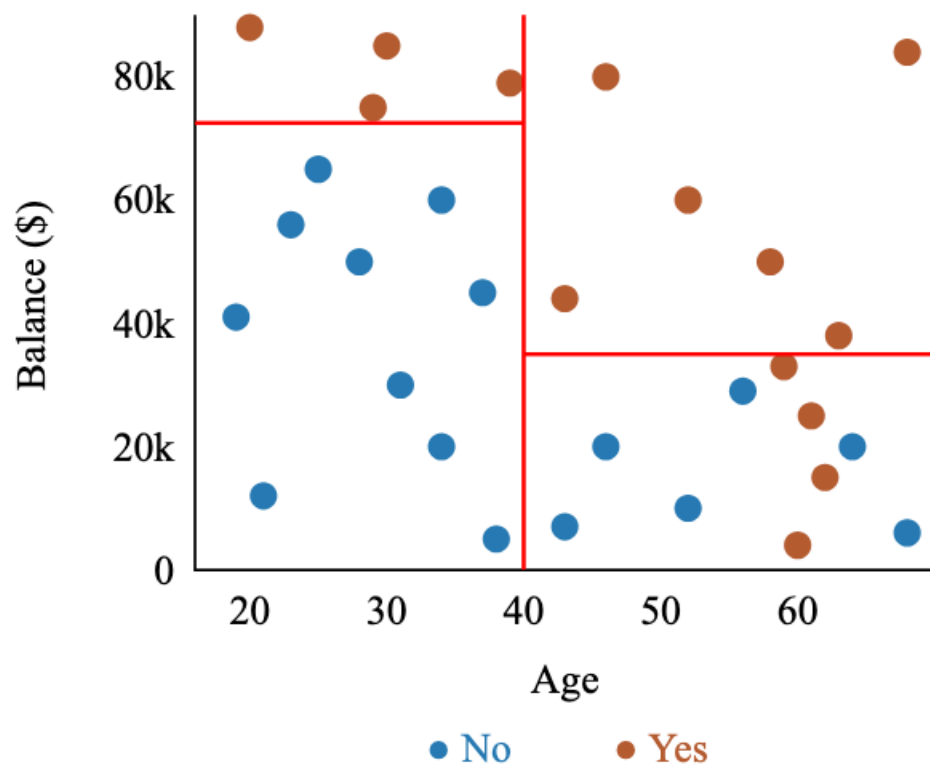
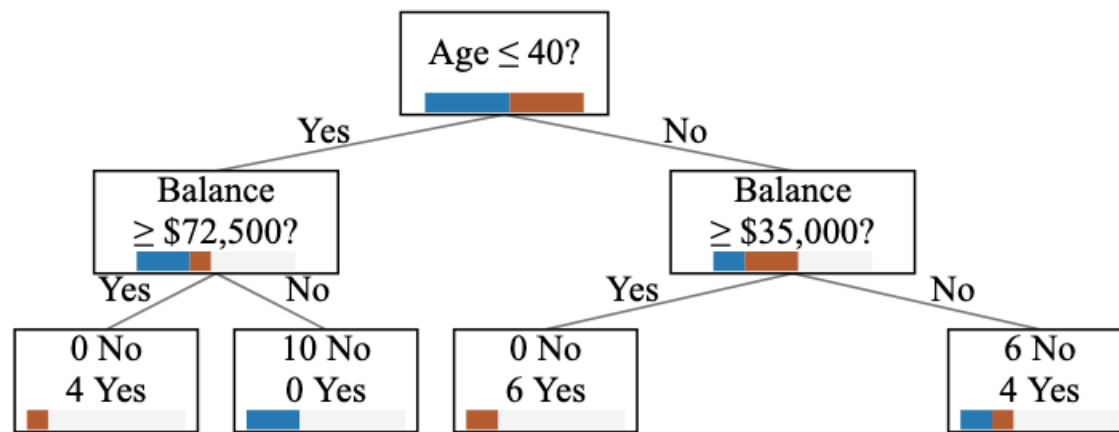
A Tree Divides Customers by Asking Questions

Ask About Balance Within the Older Group



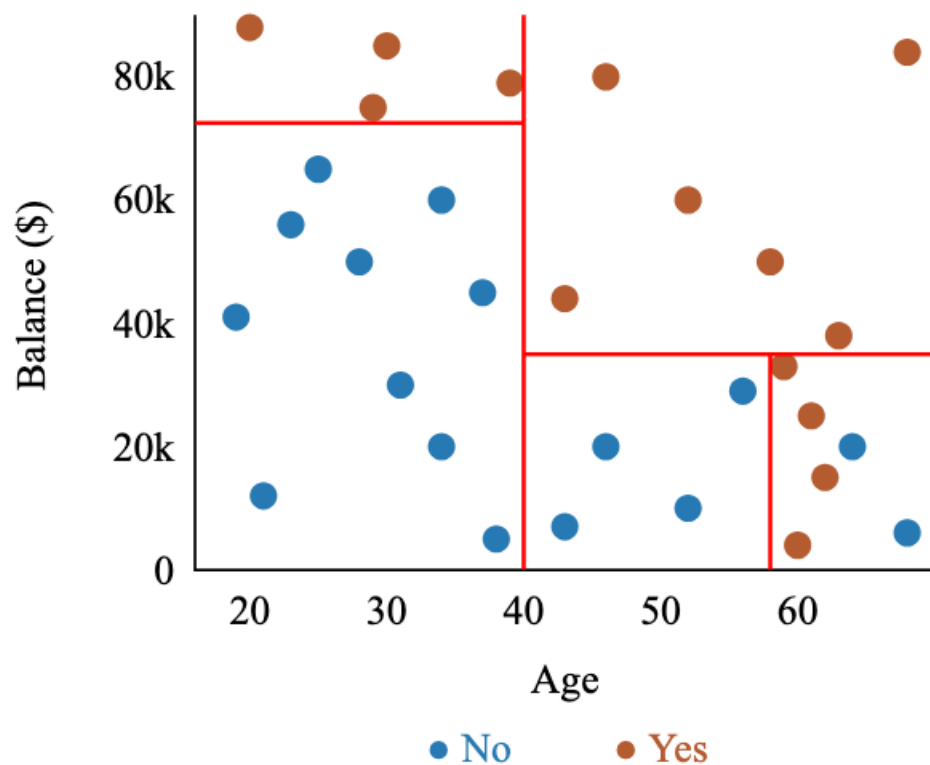
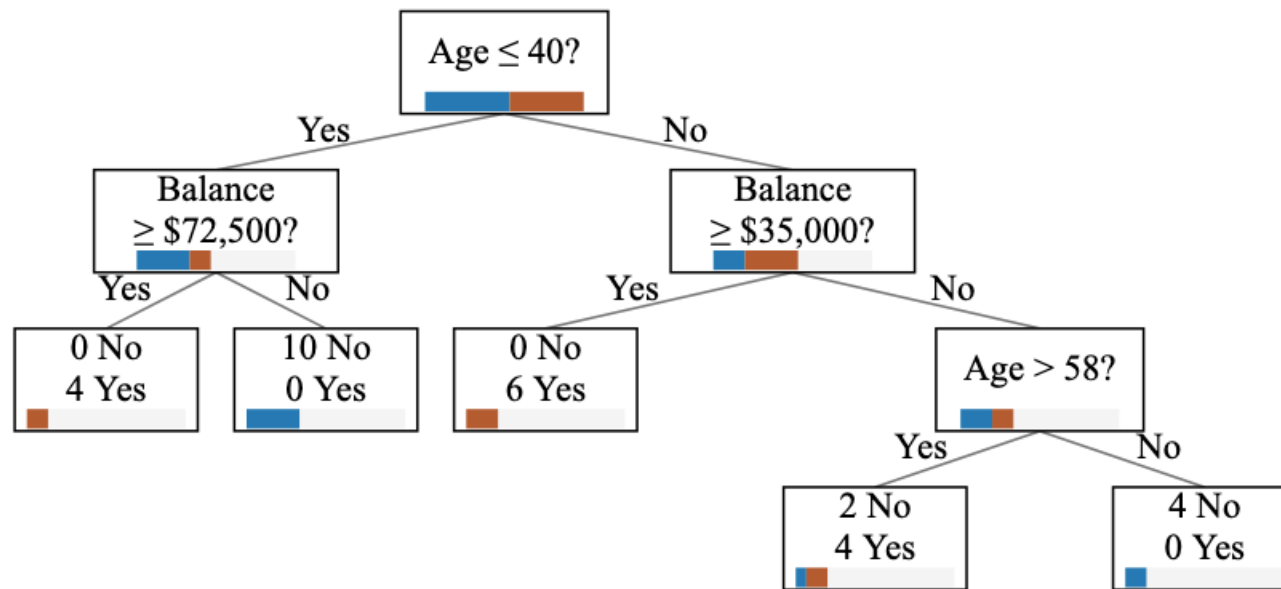
A Tree Divides Customers by Asking Questions

Split the Younger Group Too



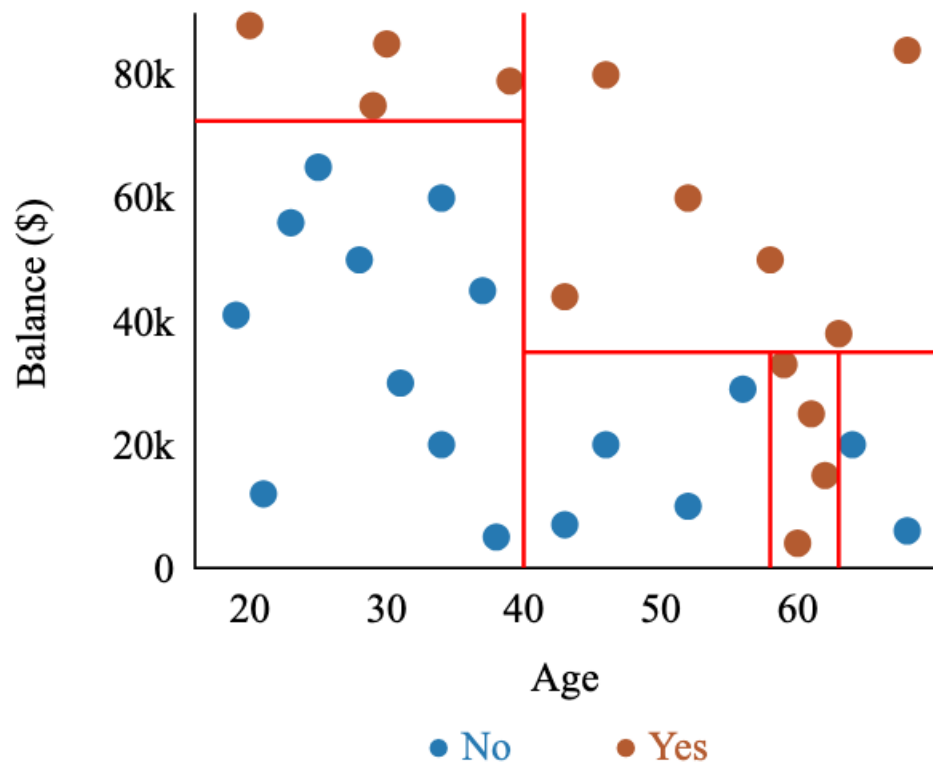
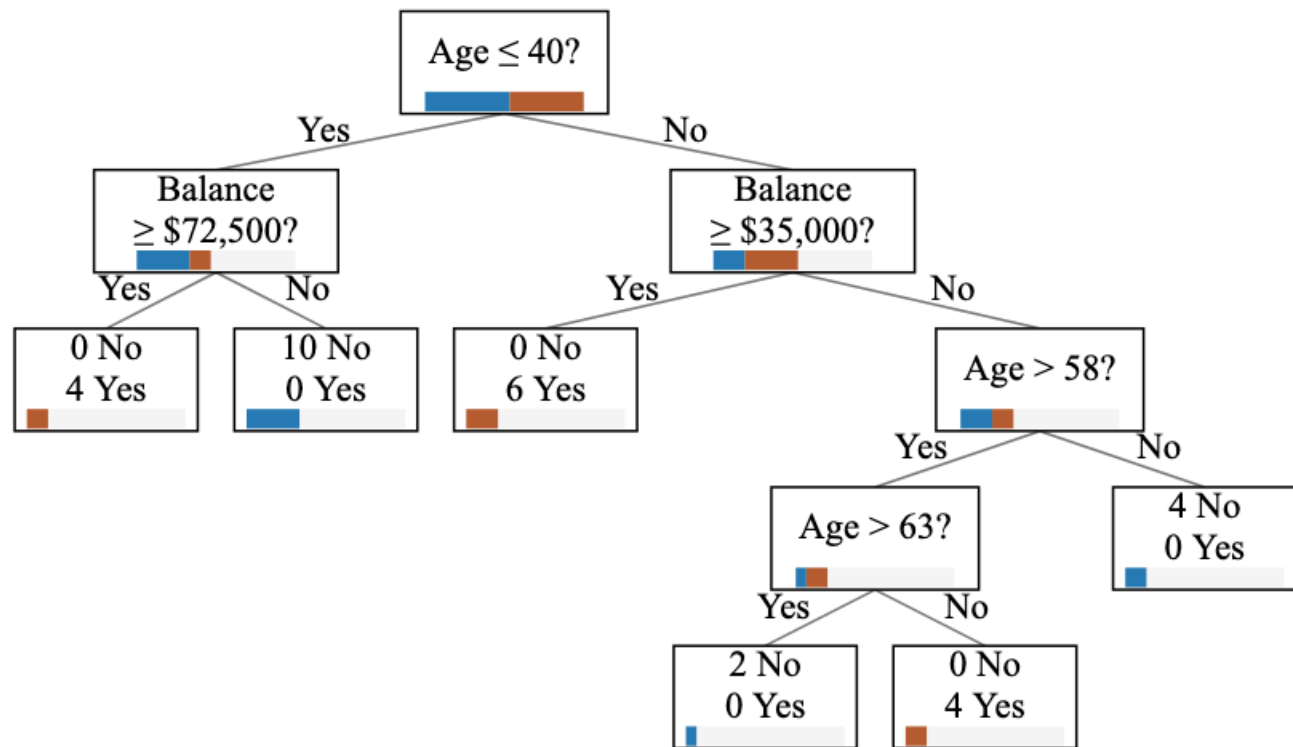
A Tree Divides Customers by Asking Questions

Ask Again Within the Older, Lower-Balance Group



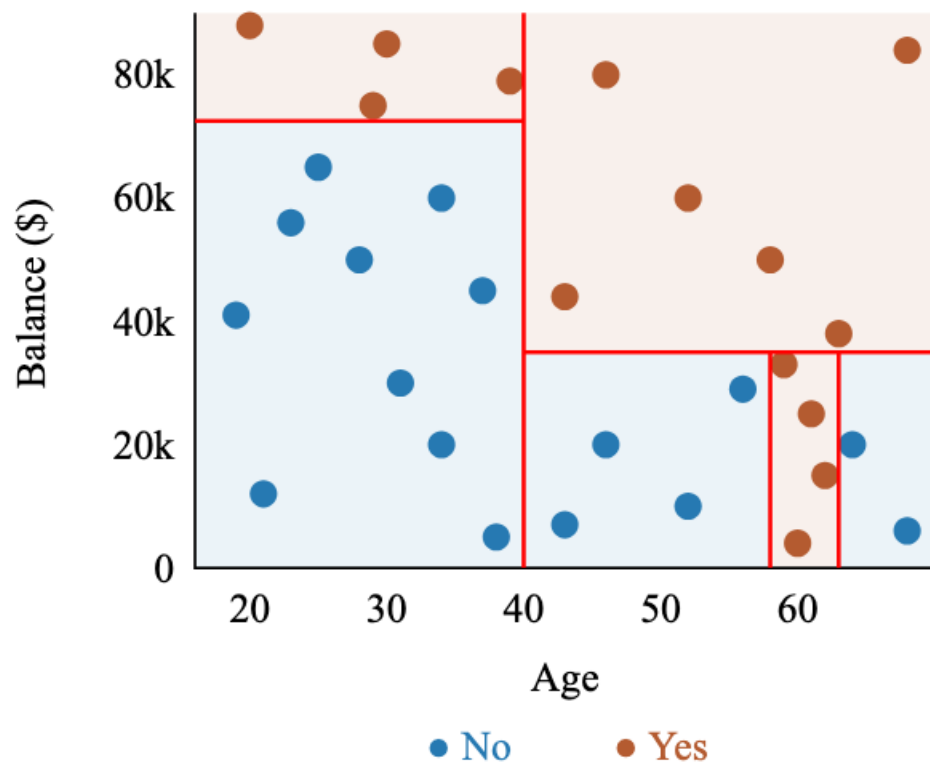
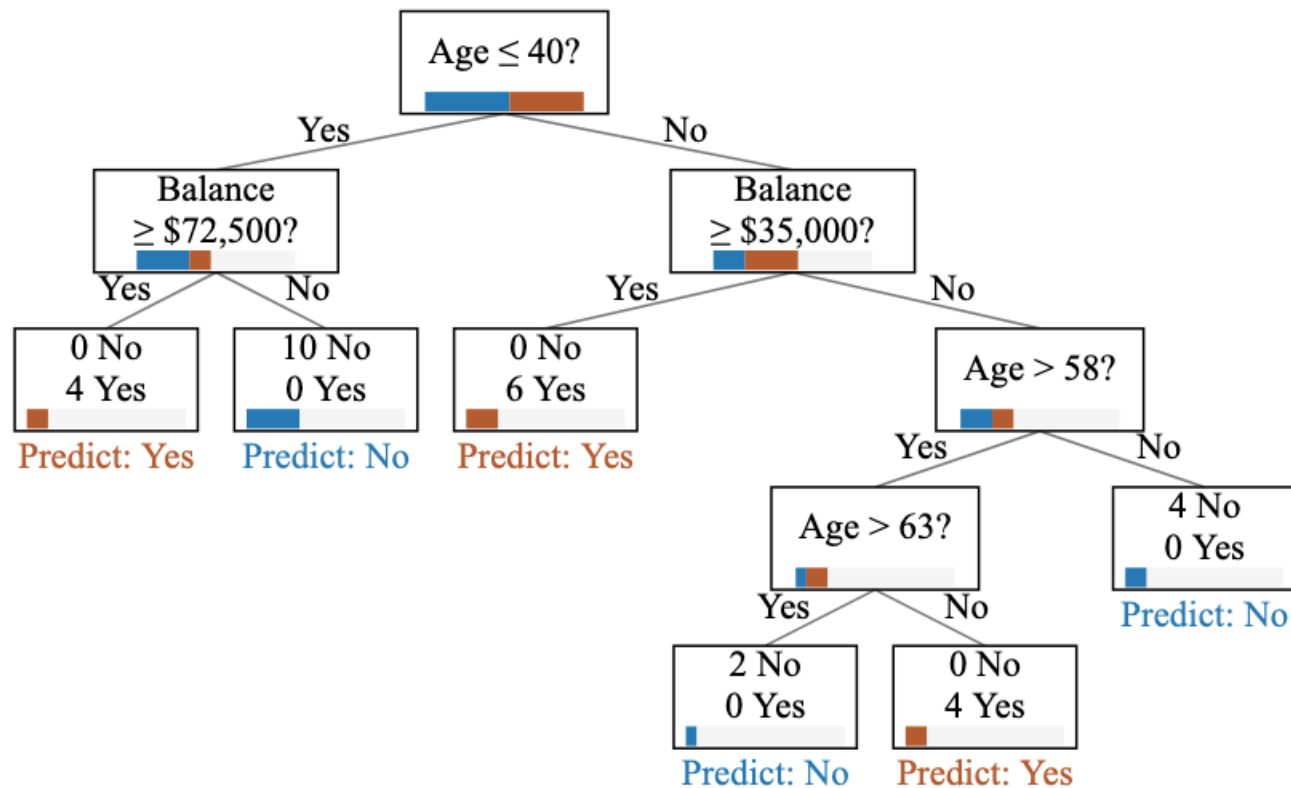
A Tree Divides Customers by Asking Questions

One More Question Within the Oldest Group



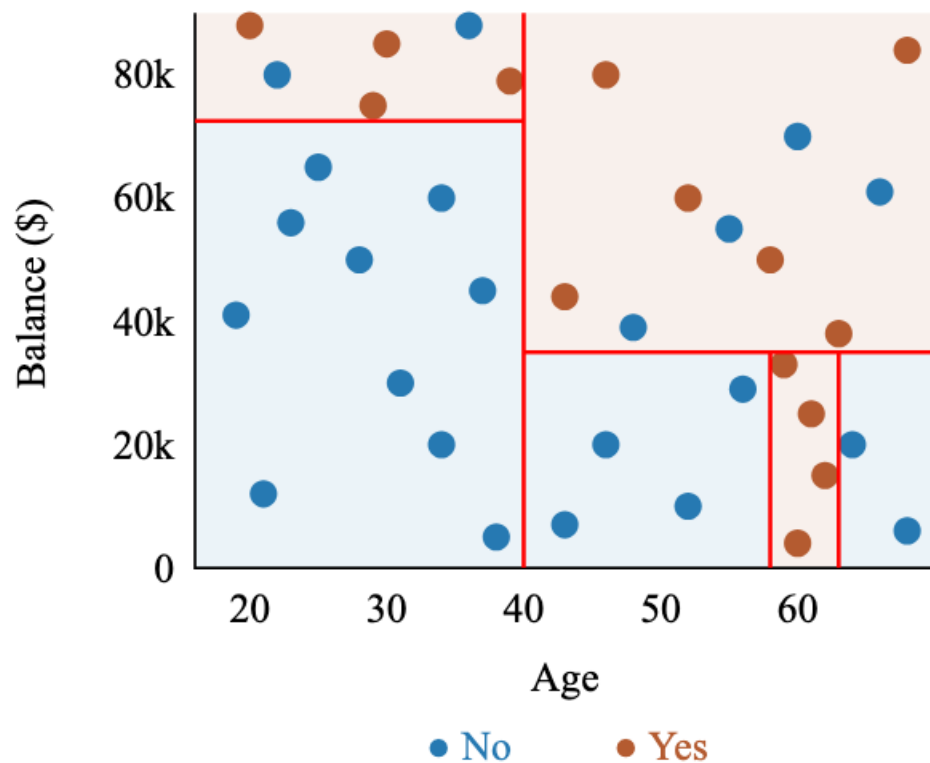
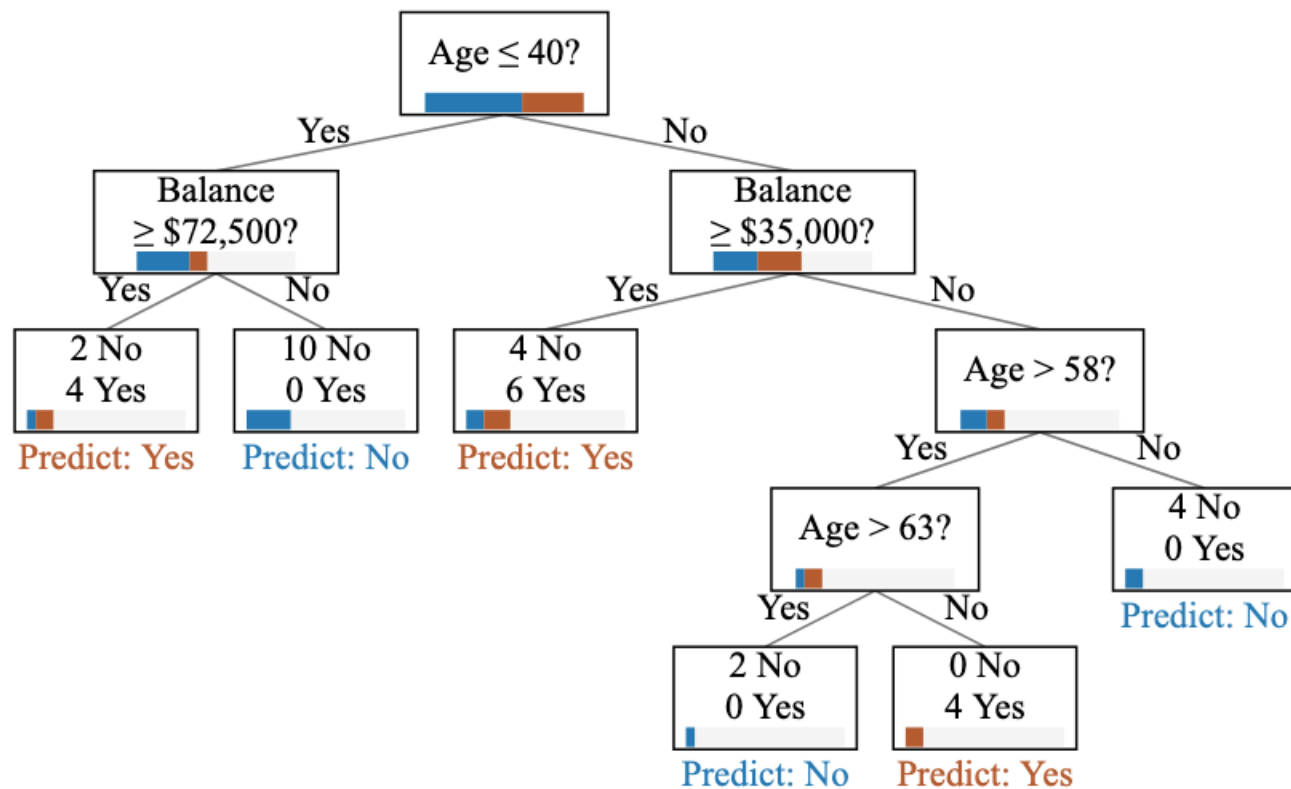
A Tree Divides Customers by Asking Questions

Every Leaf Is Pure: One Class per Group



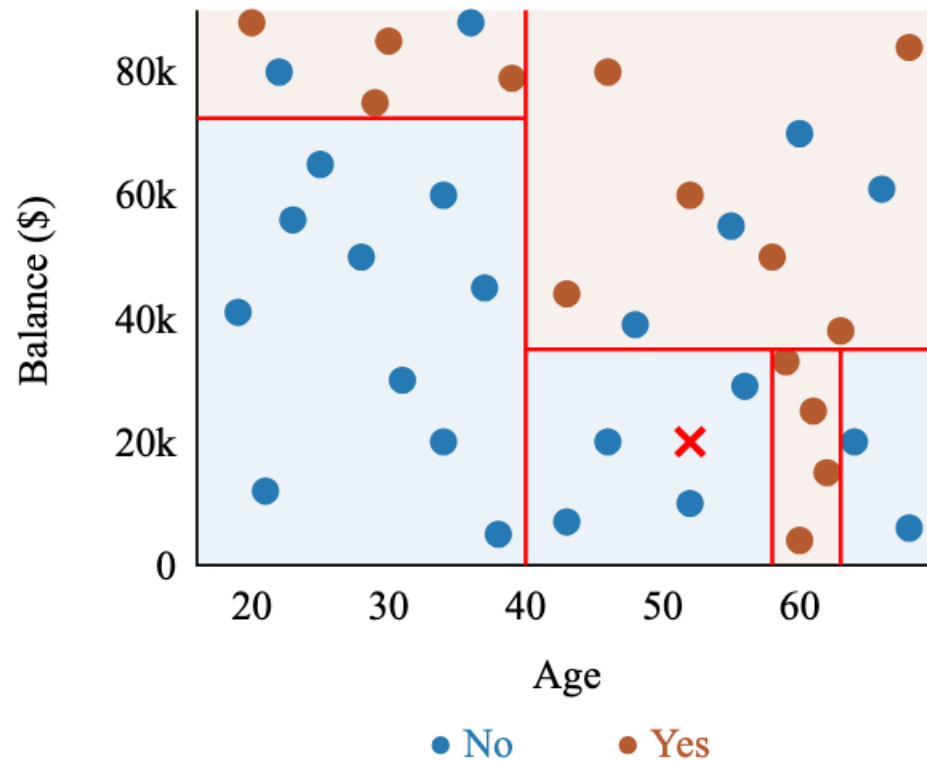
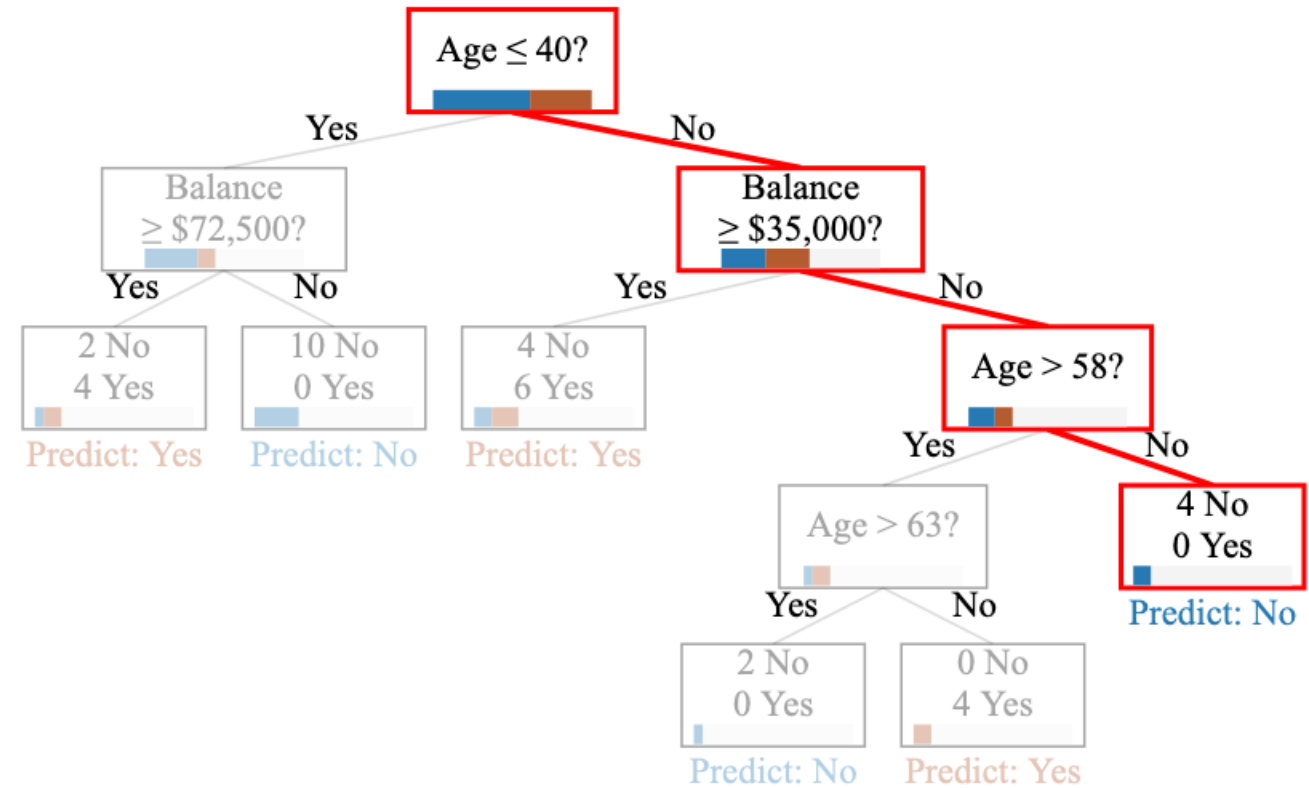
A Tree Divides Customers by Asking Questions

Example 2: All 36 Customers · Mixed Leaves Predict the Majority



Follow a Customer Through a Deeper Tree

Move the Mouse to Move the ×



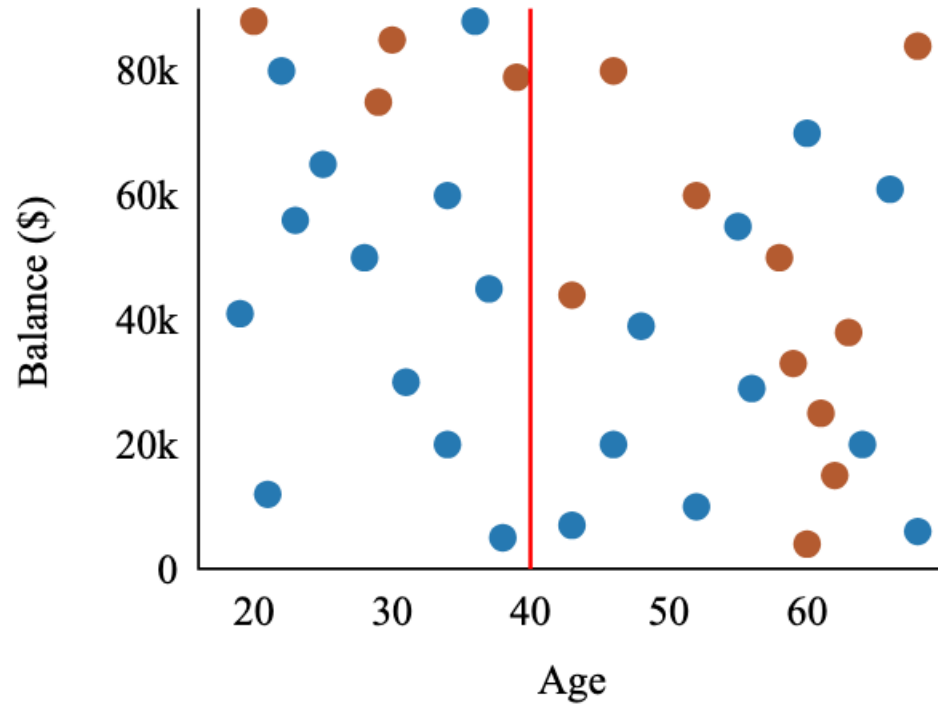
Age 52 · Balance \$20,000 · Predict: No

4

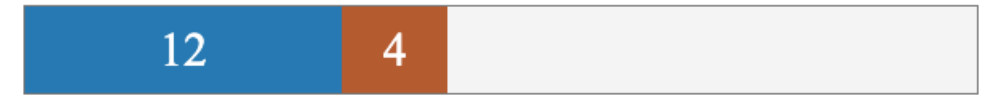
Leaf: 4 No / 0 Yes

Which Question to Ask First?

Age ≤ 40 ?



Age ≤ 40



12 No · 4 Yes

Age > 40



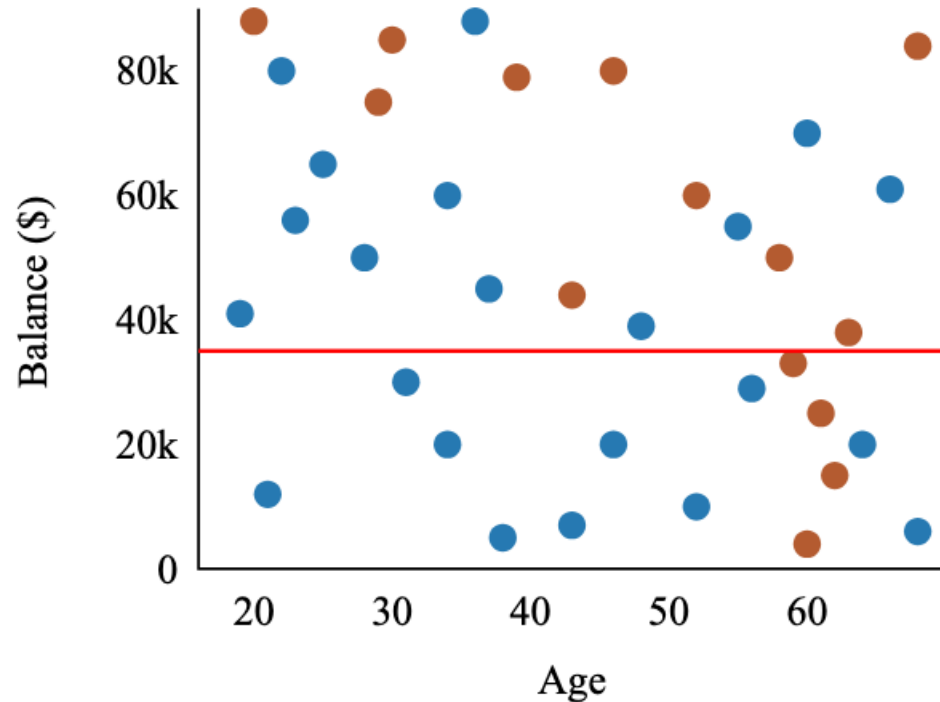
10 No · 10 Yes

● No ● Yes

Scan Age

Which Question to Ask First?

Balance $\geq$ \$35,000?

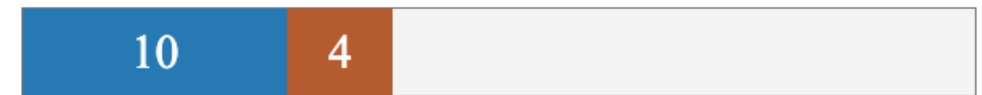


Balance $\geq$ \$35,000



12 No · 10 Yes

Balance $<$ \$35,000



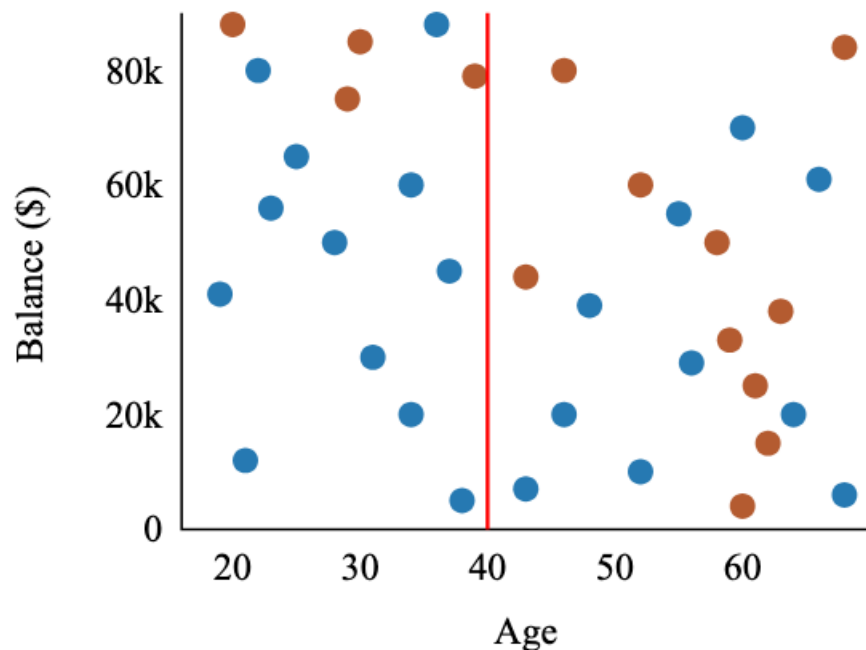
10 No · 4 Yes

• No • Yes

Scan Balance

Which Question Is Best?

Age ≤ 40 ?



Age ≤ 40



12 No · 4 Yes

Age > 40



10 No · 10 Yes

• No • Yes

Age ≤ 30 ?

Age ≤ 40 ?

Age ≤ 50 ?

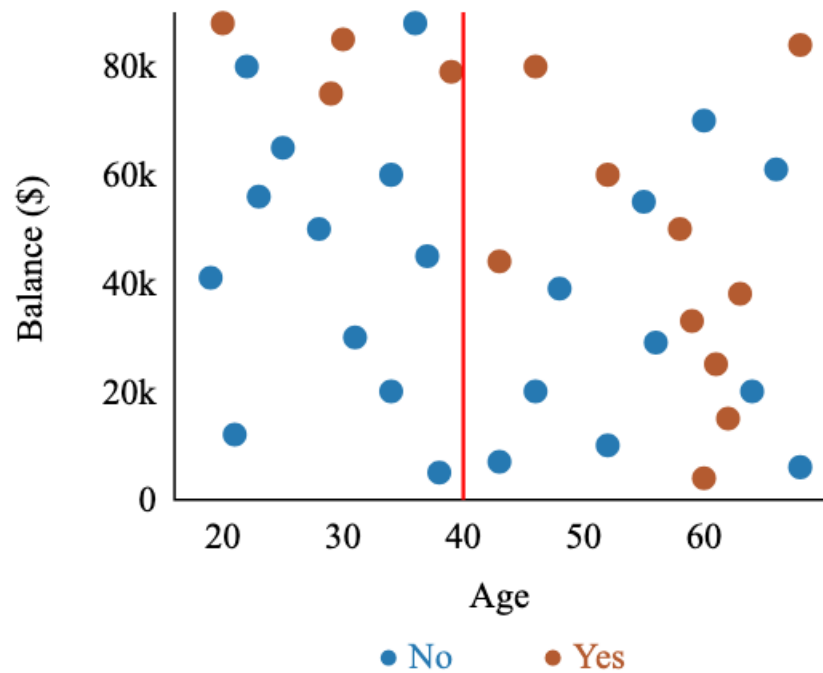
Balance $\geq \$20,000$?

Balance $\geq \$35,000$?

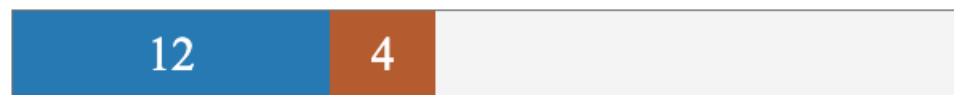
Balance $\geq \$60,000$?

From Counts to Class Fractions

Age ≤ 40 ?



Age ≤ 40



12 No · 4 Yes

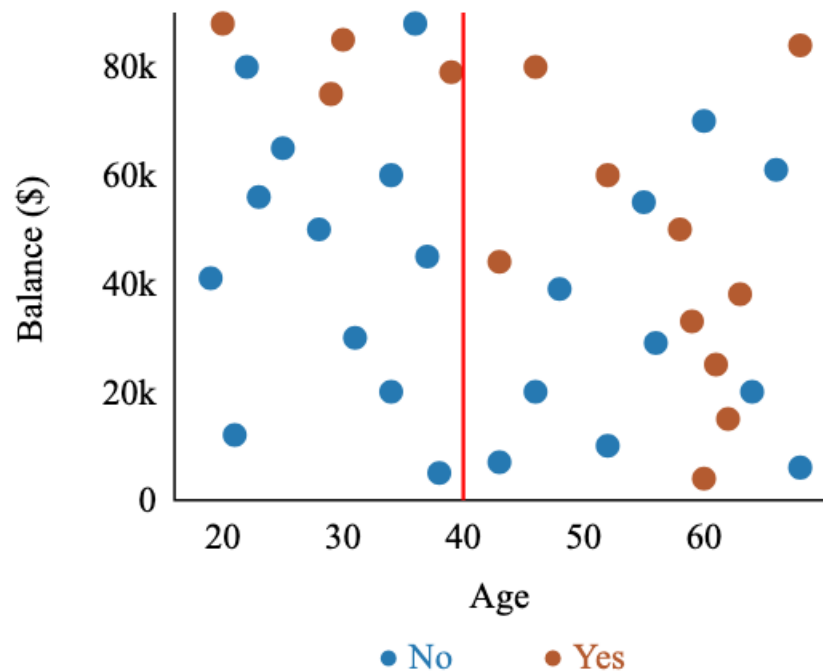
Age > 40



10 No · 10 Yes

From Counts to Class Fractions

Age ≤ 40 ?



Age ≤ 40



$$12 \text{ No} \cdot 4 \text{ Yes} \quad p(\text{No}) = \frac{12}{16} \quad p(\text{Yes}) = \frac{4}{16}$$

Age > 40



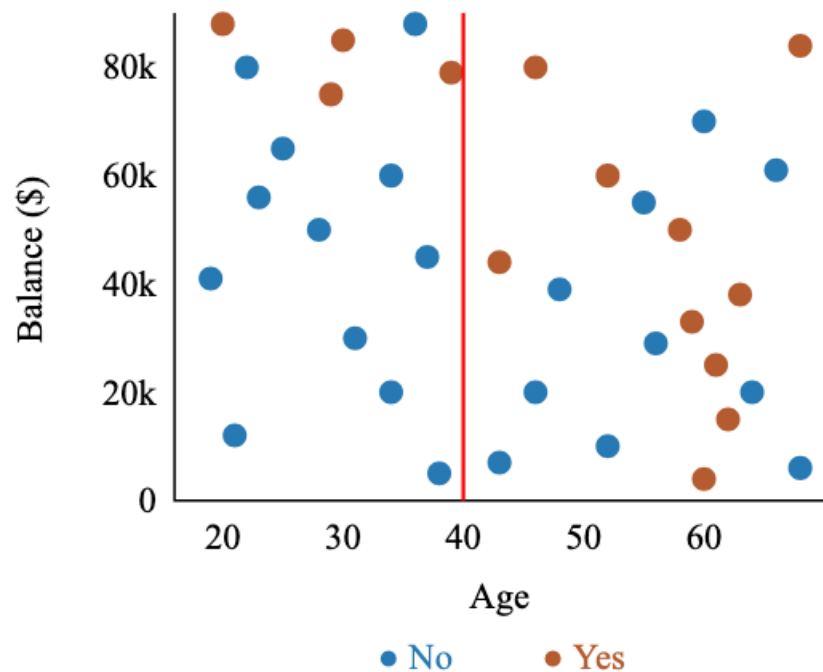
$$10 \text{ No} \cdot 10 \text{ Yes} \quad p(\text{No}) = \frac{10}{20} \quad p(\text{Yes}) = \frac{10}{20}$$

$p(\text{No})$: Fraction of the Group Labelled No

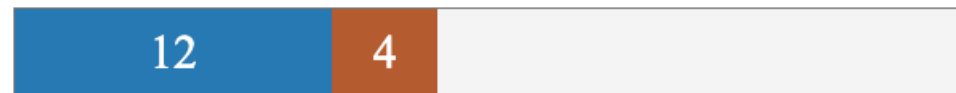
$p(\text{Yes})$: Fraction of the Group Labelled Yes

How Mixed Are the Groups?

Age ≤ 40 ?



Age ≤ 40



$$12 \text{ No} \cdot 4 \text{ Yes} \quad p(\text{No}) = \frac{12}{16} \quad p(\text{Yes}) = \frac{4}{16}$$

Age > 40



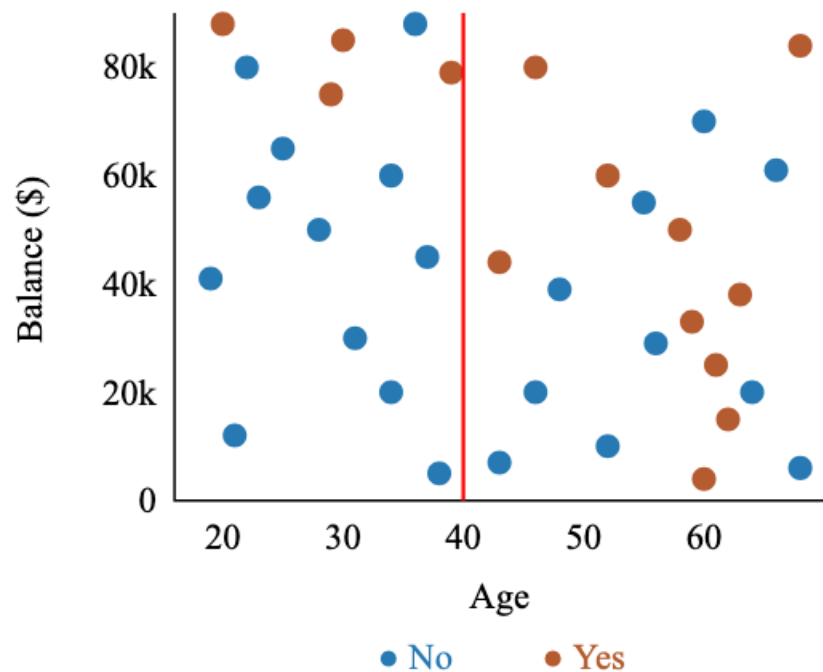
$$10 \text{ No} \cdot 10 \text{ Yes} \quad p(\text{No}) = \frac{10}{20} \quad p(\text{Yes}) = \frac{10}{20}$$

$$G = 1 - p(\text{No})^2 - p(\text{Yes})^2$$

0: Pure · 0.5: Even Yes/No Mix

How Mixed Are the Groups?

Age ≤ 40 ?



Age ≤ 40



$$12 \text{ No} \cdot 4 \text{ Yes} \quad p(\text{No}) = \frac{12}{16} \quad p(\text{Yes}) = \frac{4}{16}$$

Gini = 0.375

Age > 40



$$10 \text{ No} \cdot 10 \text{ Yes} \quad p(\text{No}) = \frac{10}{20} \quad p(\text{Yes}) = \frac{10}{20}$$

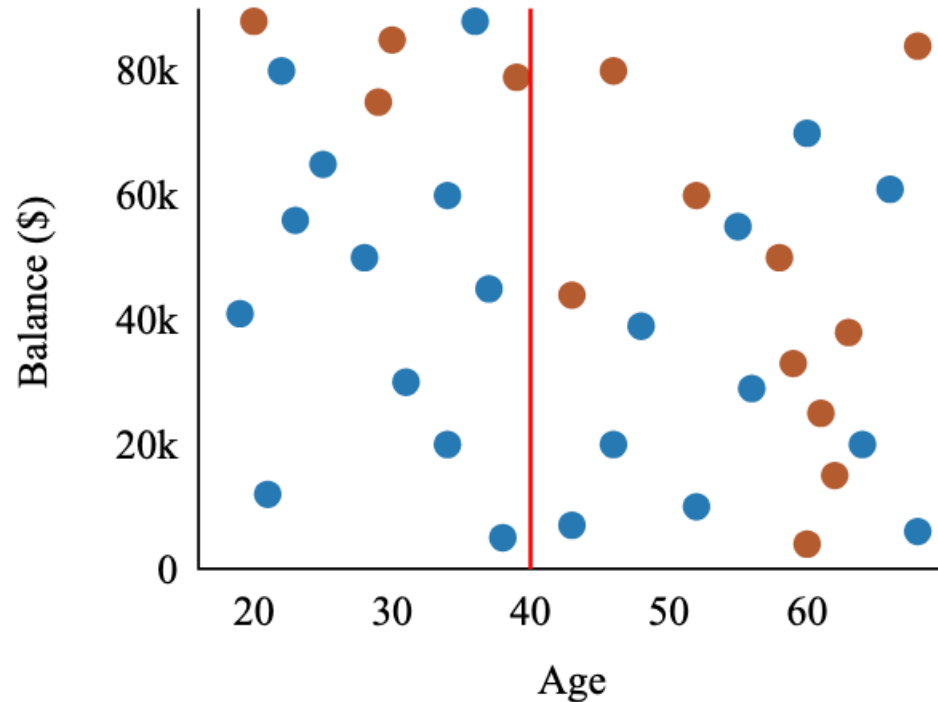
Gini = 0.500

$$G = 1 - p(\text{No})^2 - p(\text{Yes})^2$$

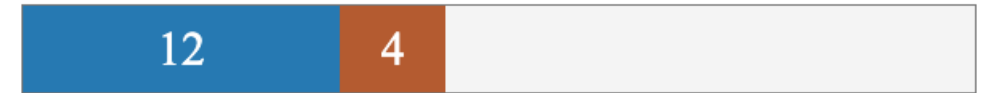
0: Pure · 0.5: Even Yes/No Mix

Which Question to Ask First?

Age ≤ 40 ?



Age ≤ 40



12 No · 4 Yes

Gini = 0.375

Age > 40



10 No · 10 Yes

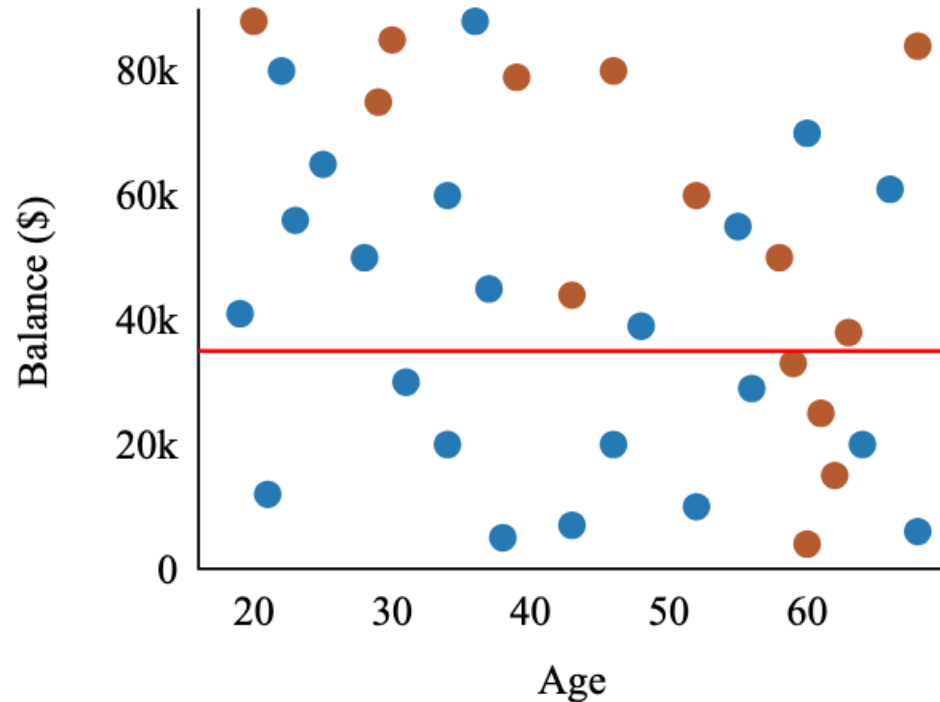
Gini = 0.500

● No ● Yes

Scan Age

Which Question to Ask First?

Balance $\geq$ \$35,000?



• No • Yes

Balance $\geq$ \$35,000



12 No · 10 Yes

Gini = 0.496

Balance < \$35,000



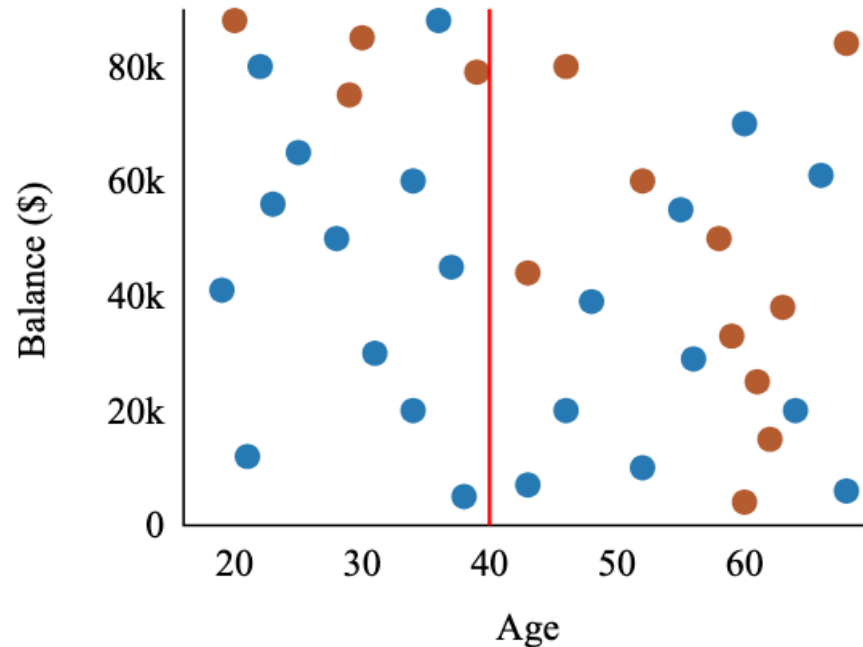
10 No · 4 Yes

Gini = 0.408

Scan Balance

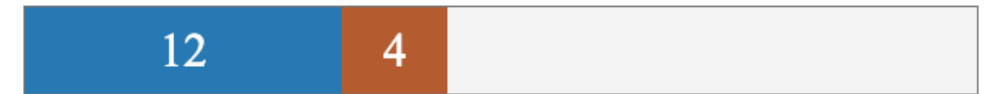
How Does Gini Help Us?

Age ≤ 40 ?



Original Gini = 0.475

Age ≤ 40



12 No · 4 Yes

Gini = 0.375

Age > 40



10 No · 10 Yes

Gini = 0.500

● No ● Yes

Age ≤ 30 ?

Age ≤ 40 ?

Age ≤ 50 ?

Balance $\geq \$20,000$?

Balance $\geq \$35,000$?

Balance $\geq \$60,000$?

What Makes a Good Split?

How Much Less Mixed Are the Groups After the Split?

What Makes a Good Split?

How Much Less Mixed Are the Groups After the Split?

Gini Before – Weighted Gini After

What Makes a Good Split?

How Much Less Mixed Are the Groups After the Split?

Gini Before – Weighted Gini After

Larger Gini Reduction → Better Split

Gini Reduction

Before the Split



22 No · 14 Yes

Gini = 0.475

Gini Reduction

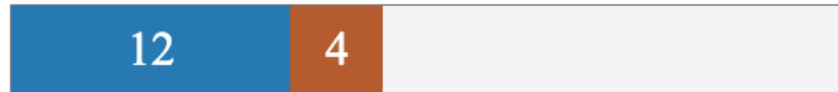
Before the Split



22 No · 14 Yes

Gini = 0.475

Age $\leq$ 40



12 No · 4 Yes

Gini = 0.375

Age $>$ 40



10 No · 10 Yes

Gini = 0.500

Gini Reduction

Before the Split



22 No · 14 Yes

Gini = 0.475

Age $\leq$ 40



12 No · 4 Yes

Gini = 0.375

Age $>$ 40



10 No · 10 Yes

Gini = 0.500

Weight Each Group by Its Share of Customers

$$(16/36) \times 0.375 + (20/36) \times 0.500 \approx 0.444$$

Gini Reduction

Before the Split



22 No · 14 Yes

Gini = 0.475

Age $\leq$ 40



12 No · 4 Yes

Gini = 0.375

Age $>$ 40



10 No · 10 Yes

Gini = 0.500

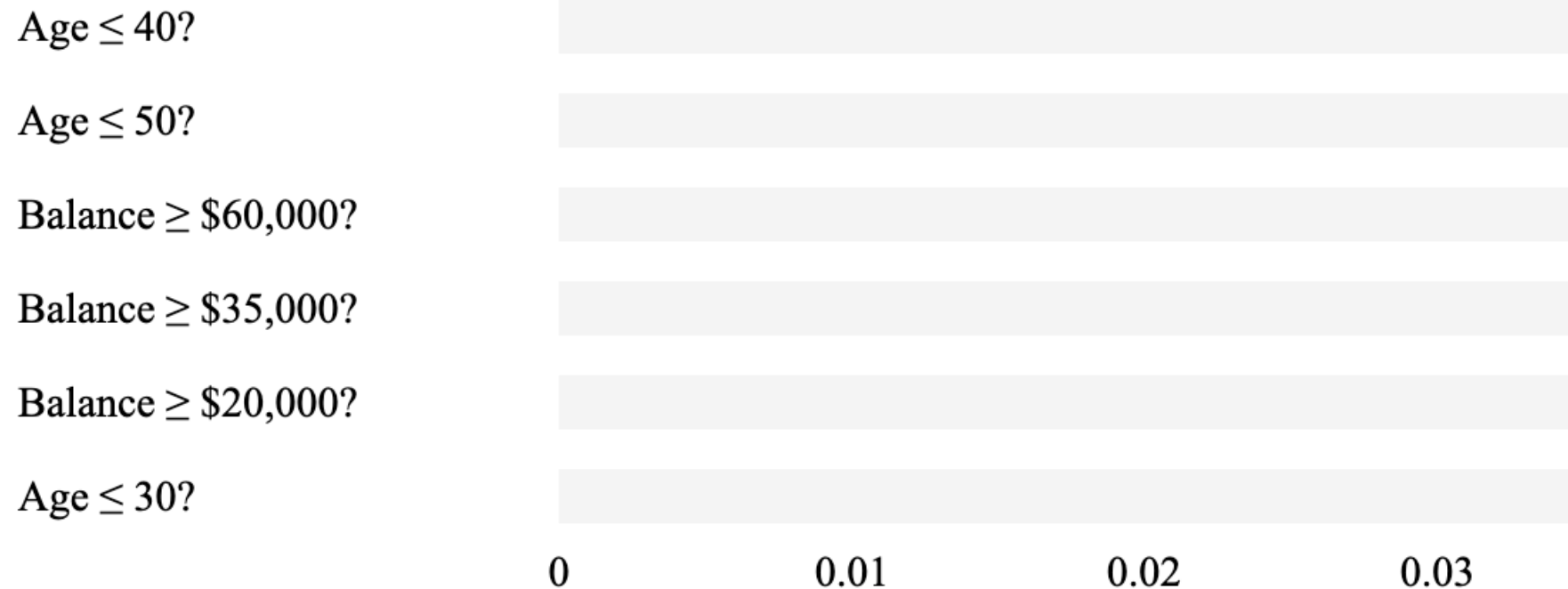
Weight Each Group by Its Share of Customers

$$(16/36) \times 0.375 + (20/36) \times 0.500 \approx 0.444$$

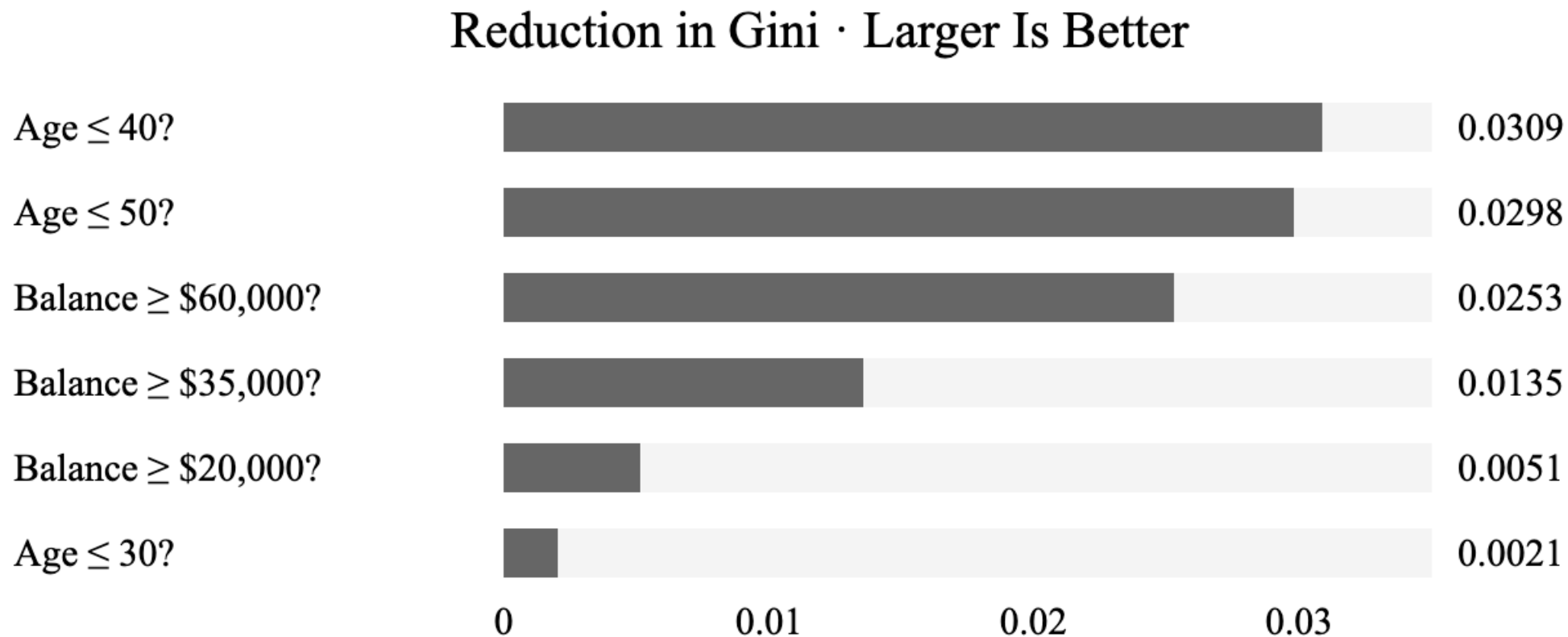
Gini Reduction: $0.475 - 0.444 \approx 0.031$

Which Question Reduces Impurity Most?

Reduction in Gini · Larger Is Better

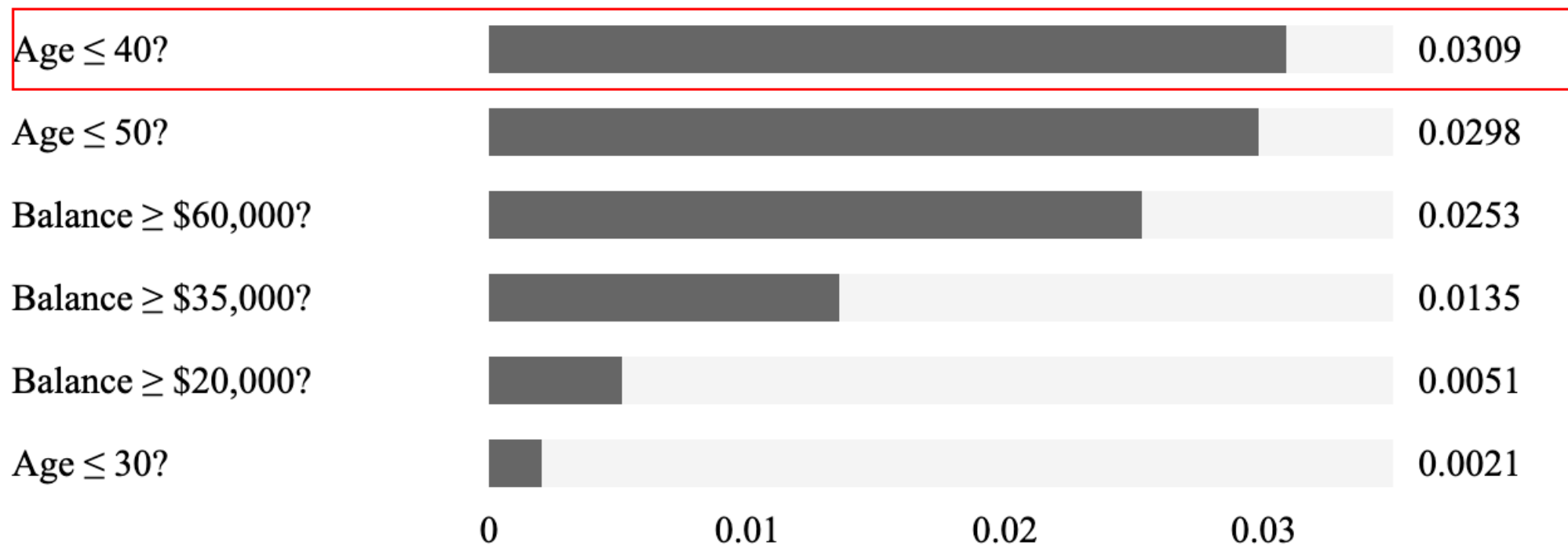


Which Question Reduces Impurity Most?



Which Question Reduces Impurity Most?

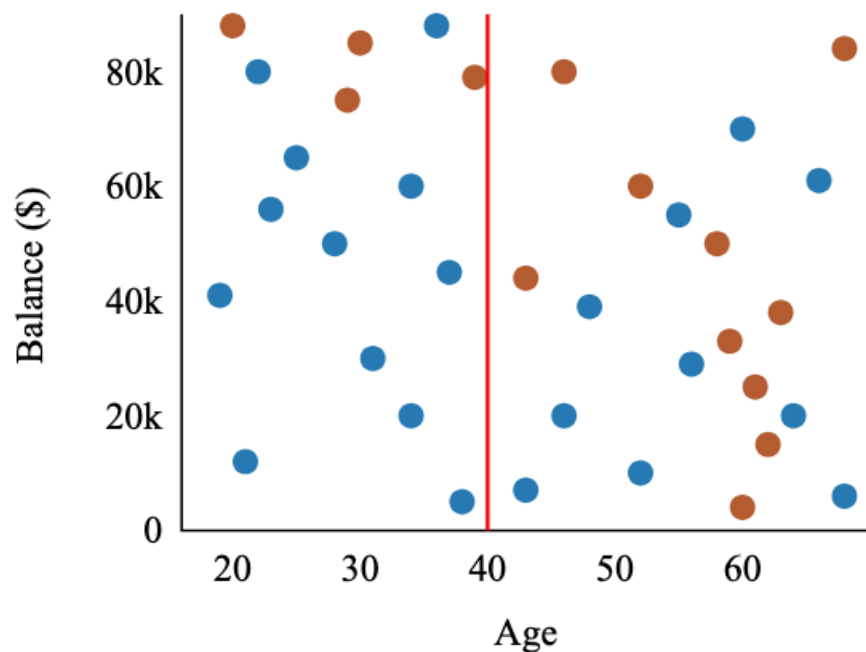
Reduction in Gini · Larger Is Better



Best of These Six: Age ≤ 40?

Find the Split With the Largest Gini Reduction

Age ≤ 40 ?



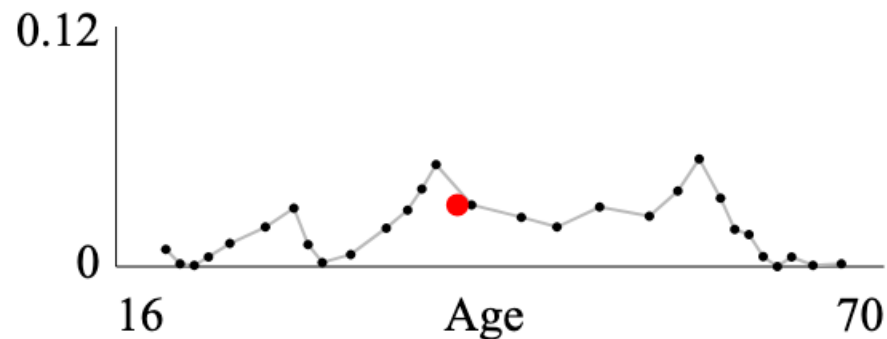
Age ≤ 40



Age > 40



Gini Reduction: 0.0309

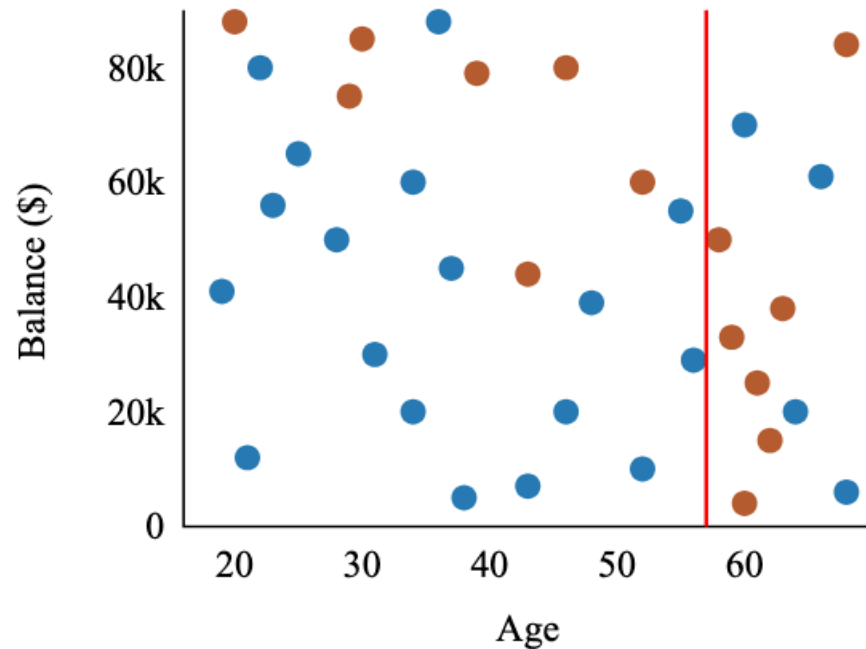


● No ● Yes

Scan Age

Find the Split With the Largest Gini Reduction

Age ≤ 57 ?



● No ● Yes

Age ≤ 57

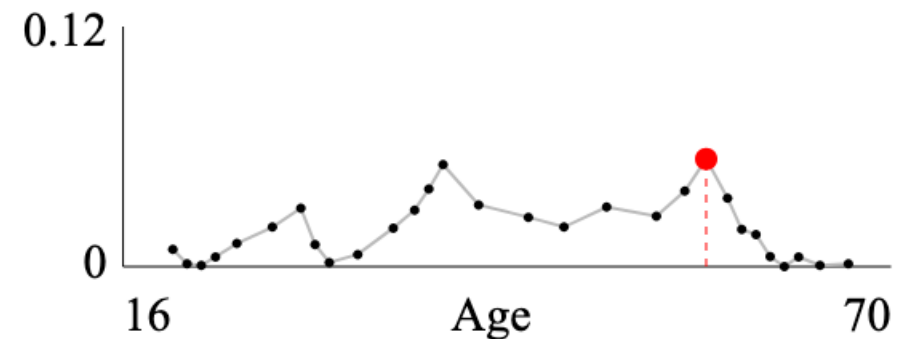


Age > 57



Gini Reduction: 0.0539

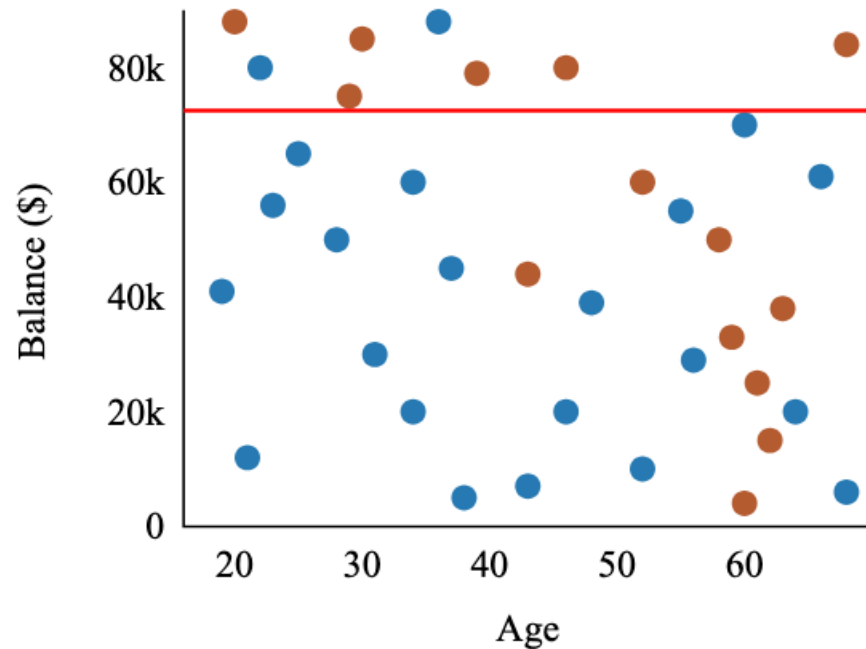
Best



Scan Age

Find the Split With the Largest Gini Reduction

Balance $\geq$ \$72,500?



● No ● Yes

Balance $\geq$ \$72,500

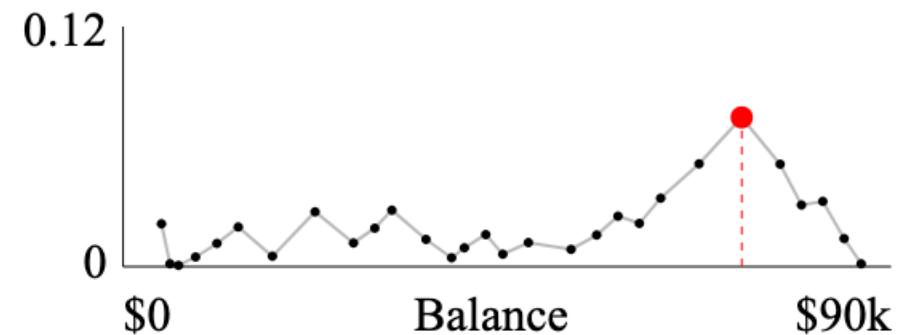


Balance $<$ \$72,500



Gini Reduction: 0.0745

Best

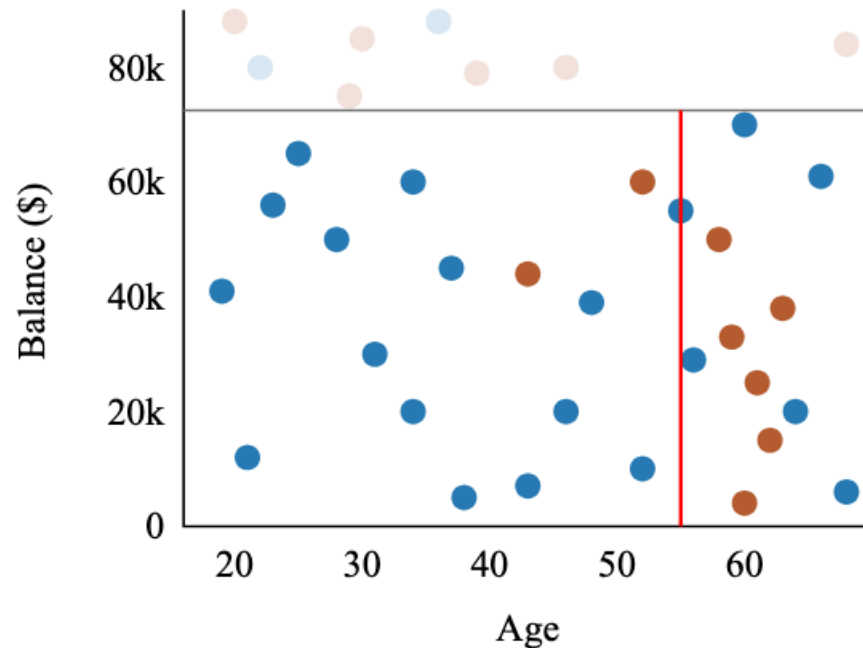


Scan Balance

Ask Again Within the Mixed Branch

Balance < \$72,500 · Starting Gini = 0.408

Age ≤ 55?



● No ● Yes

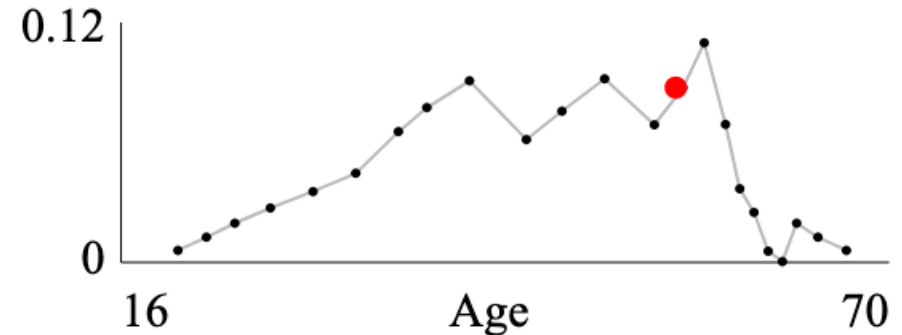
Age ≤ 55



Age > 55



Gini Reduction: 0.0873

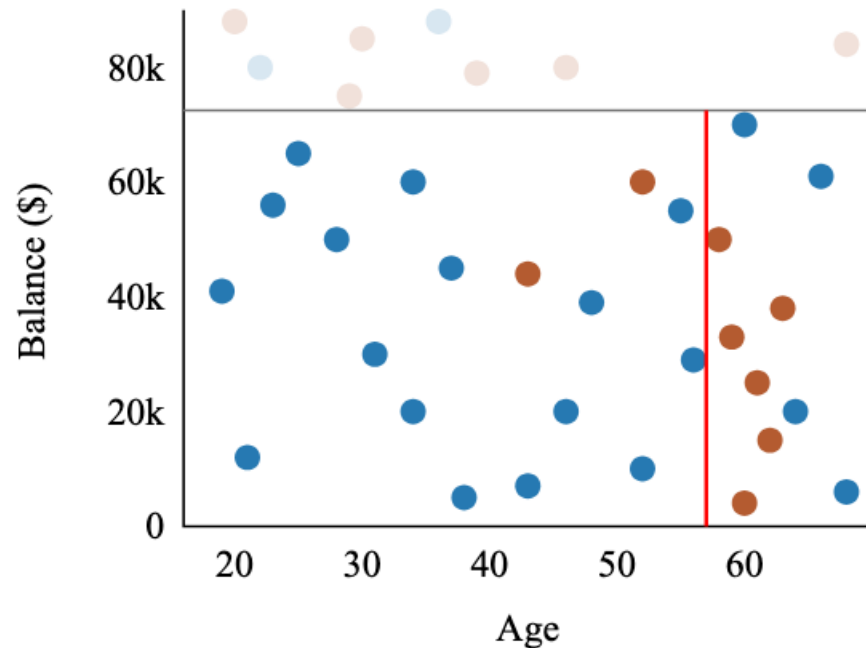


Scan Age

Ask Again Within the Mixed Branch

Balance < \$72,500 · Starting Gini = 0.408

Age ≤ 57?



● No ● Yes

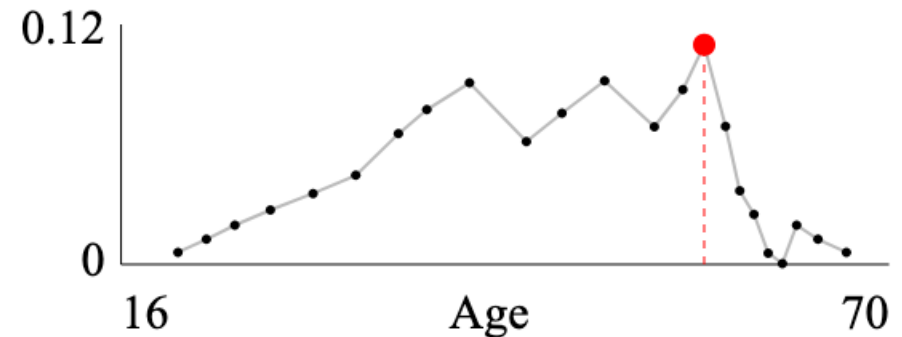
Age ≤ 57



Age > 57



Gini Reduction: 0.1098

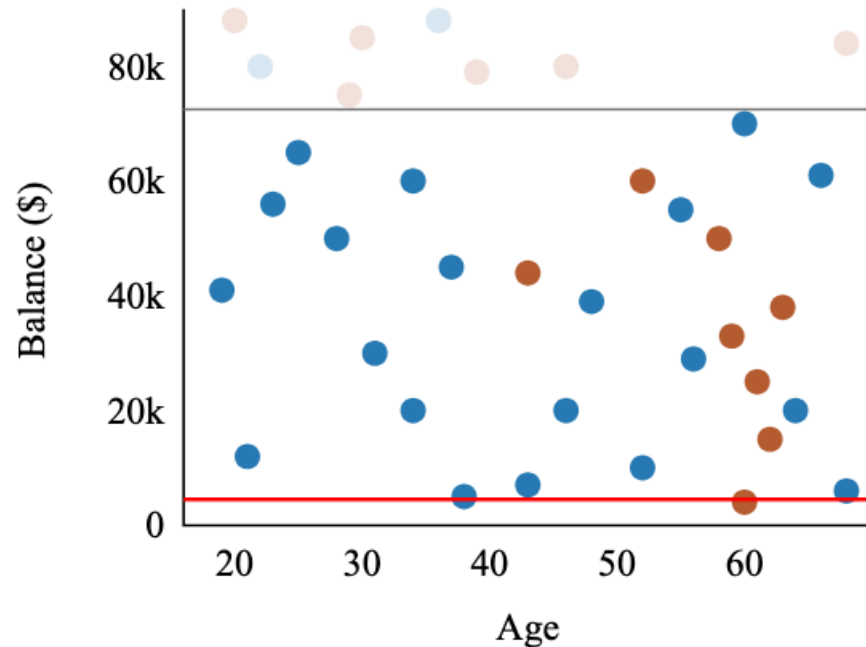


Scan Age

Ask Again Within the Mixed Branch

Balance < \$72,500 · Starting Gini = 0.408

Balance ≥ \$4,500?



● No ● Yes

Scan Balance

Balance ≥ \$4,500

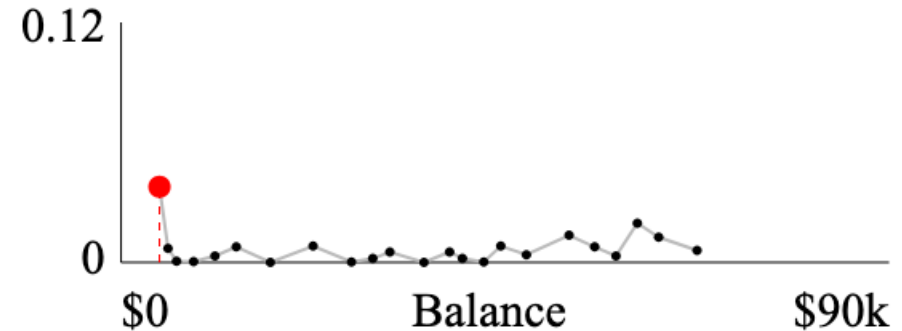


Balance < \$4,500



Gini Reduction: 0.0378

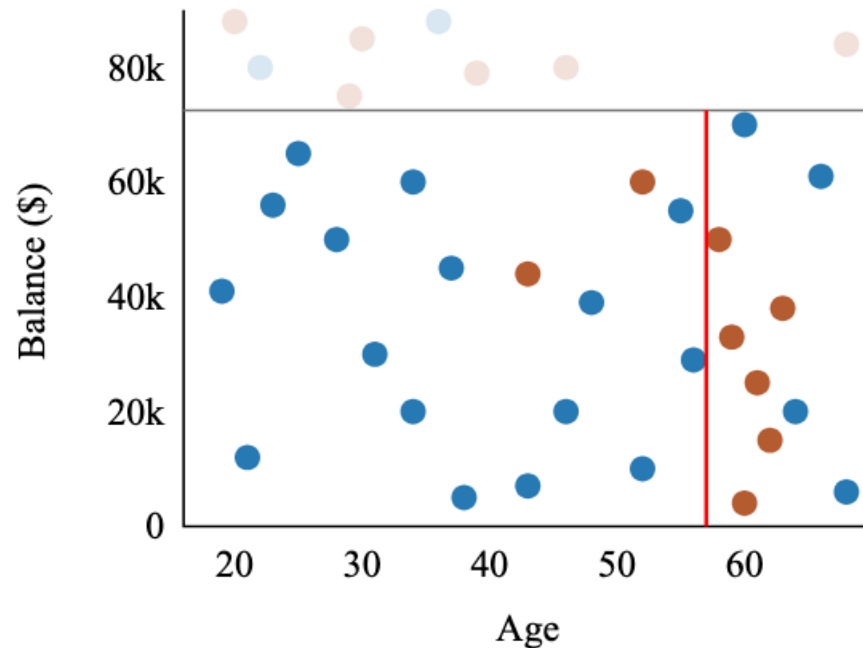
Best



Ask Again Within the Mixed Branch

Balance < \$72,500 · Starting Gini = 0.408

Age ≤ 57?



● No ● Yes

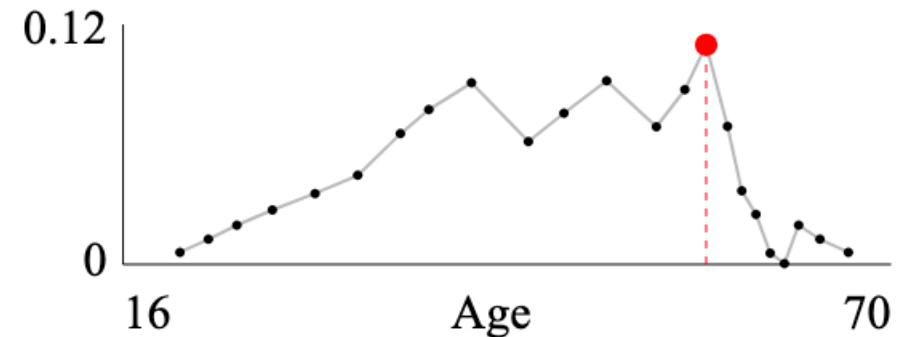
Age ≤ 57



Age > 57



Gini Reduction: 0.1098

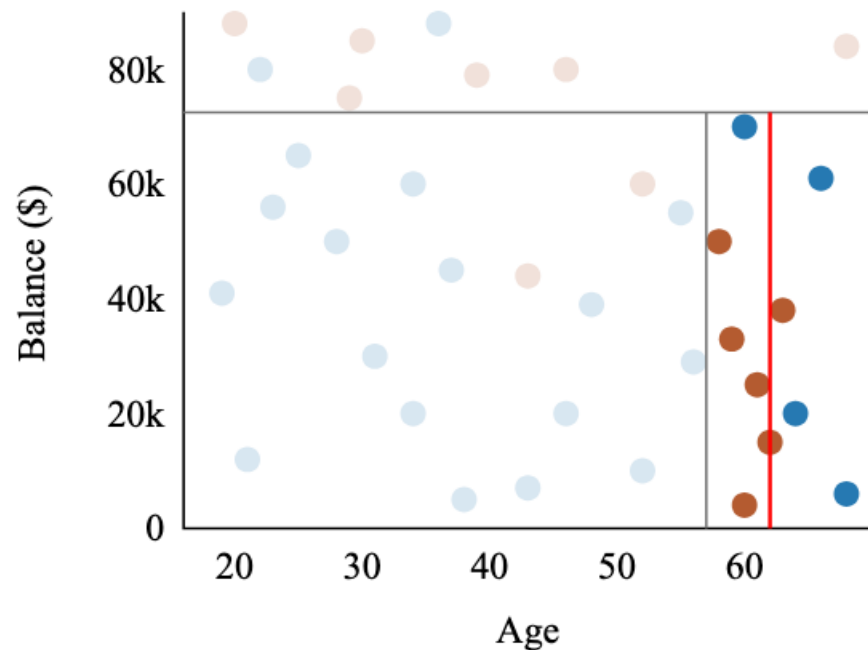


Scan Age

Ask One More Question

Balance < \$72,500 · Age > 57 · Starting Gini = 0.480

Age ≤ 62?



● No ● Yes

Age ≤ 62



Age > 62



Gini Reduction: 0.1633

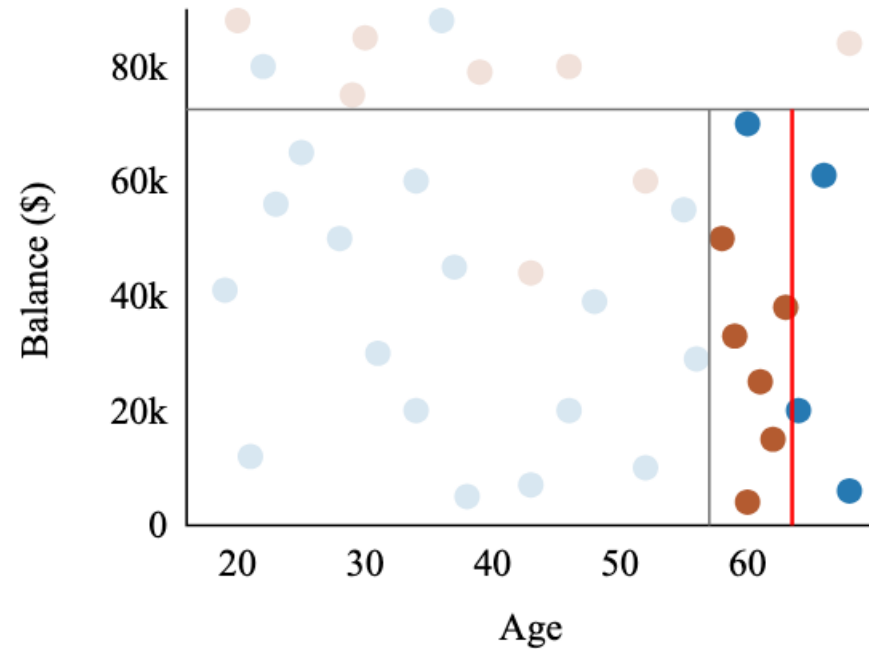


Scan Age

Ask One More Question

Balance < \$72,500 · Age > 57 · Starting Gini = 0.480

Age ≤ 63.5?



● No ● Yes

Scan Age

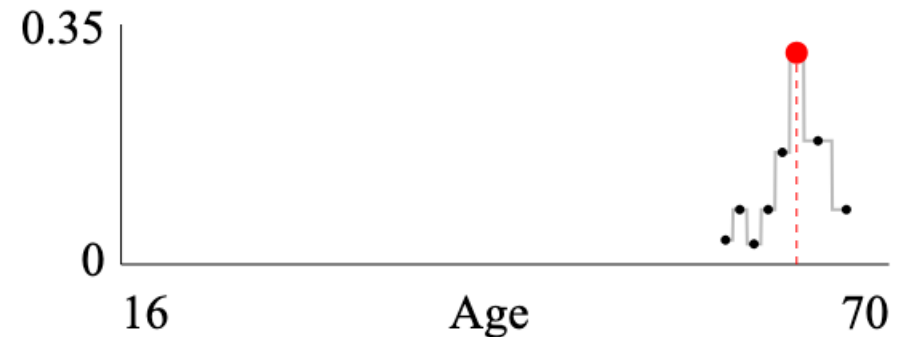
Age ≤ 63.5



Age > 63.5



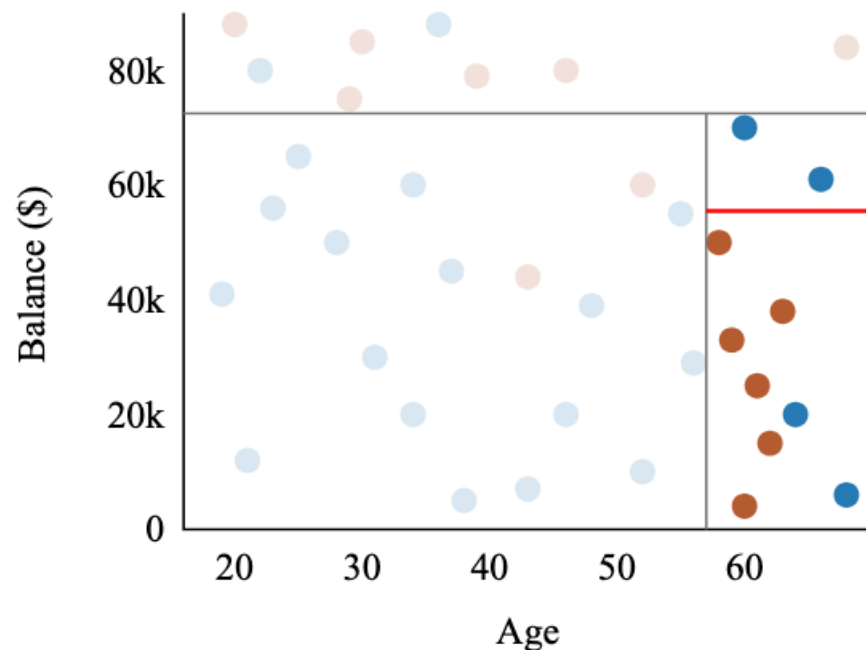
Gini Reduction: 0.3086



Ask One More Question

Balance < \$72,500 · Age > 57 · Starting Gini = 0.480

Balance ≥ \$55,500?



Balance ≥ \$55,500

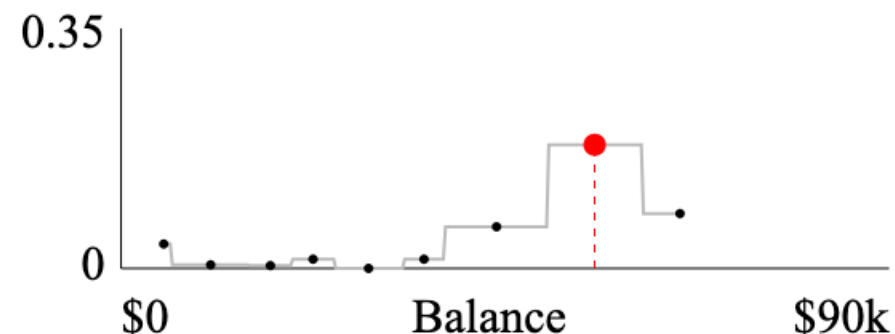


Balance < \$55,500



Gini Reduction: 0.1800

Best

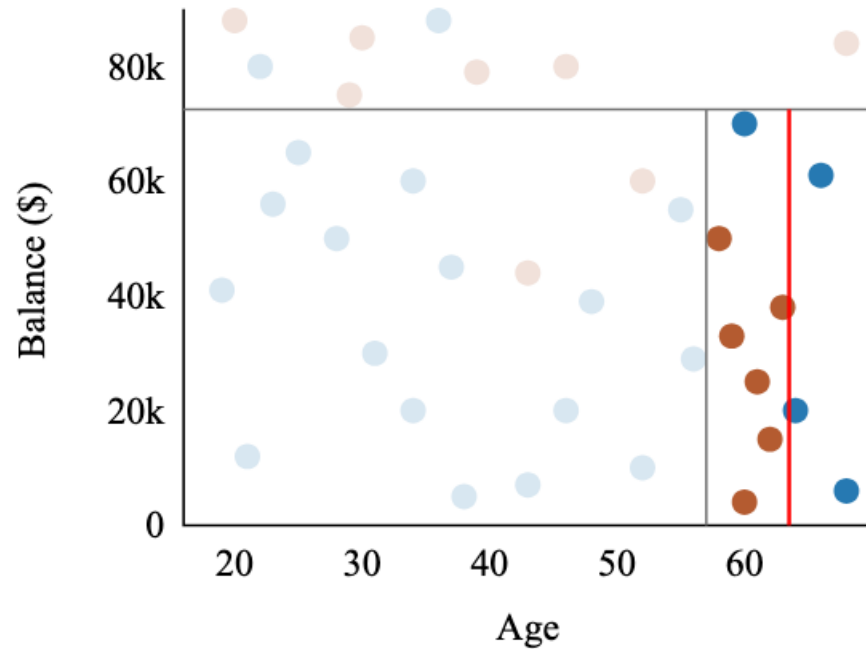


Scan Balance

Ask One More Question

Balance < \$72,500 · Age > 57 · Starting Gini = 0.480

Age ≤ 63.5?



● No ● Yes

Scan Age

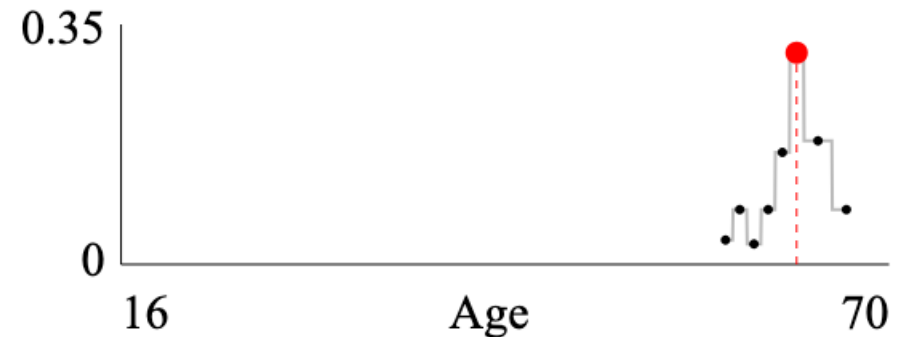
Age ≤ 63.5



Age > 63.5

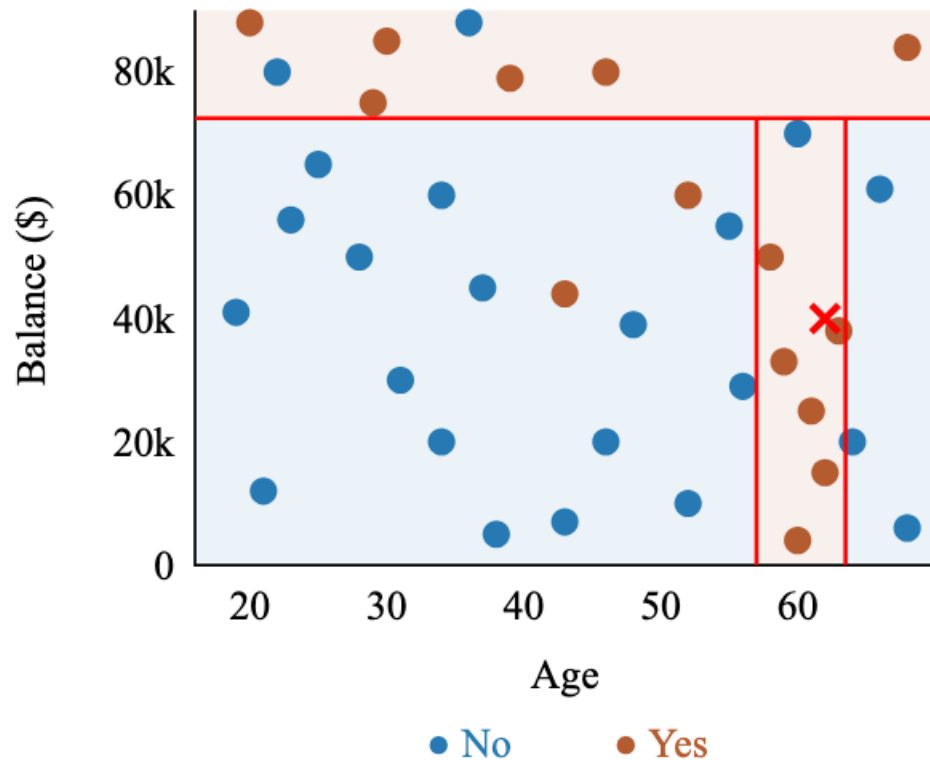


Gini Reduction: 0.3086

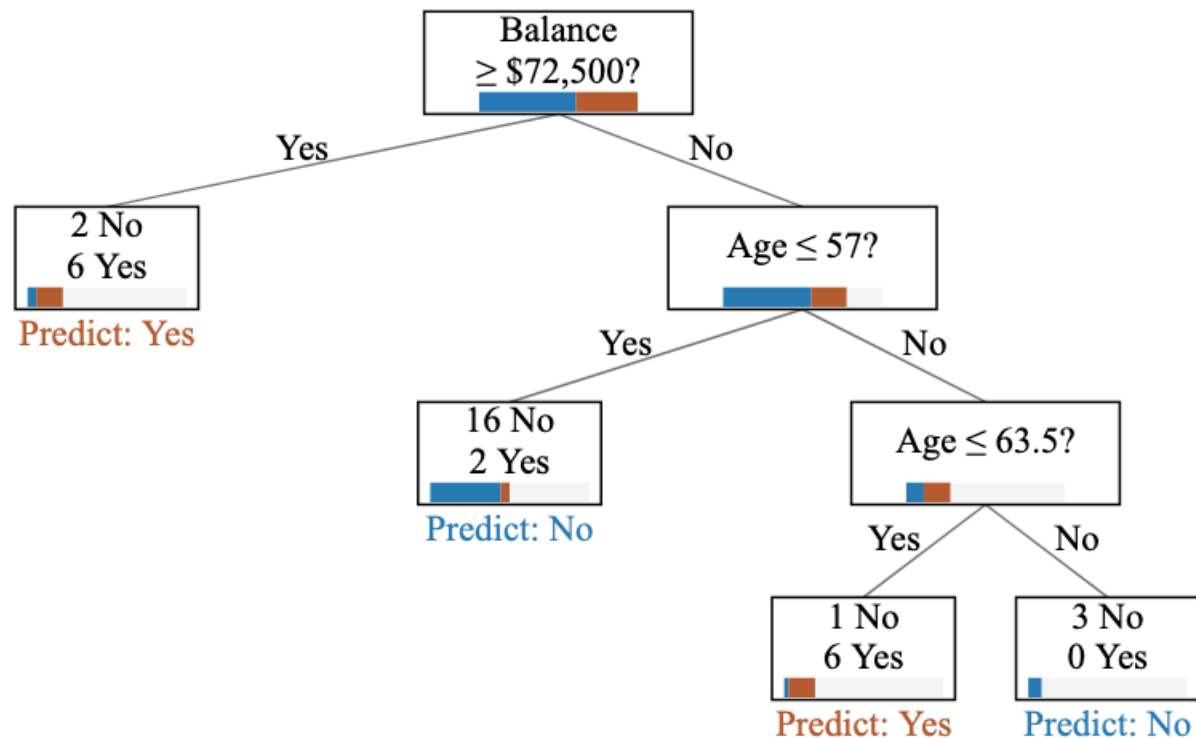


Training Builds the Rules; Prediction Uses Them

Age 62 · Balance \$40,000



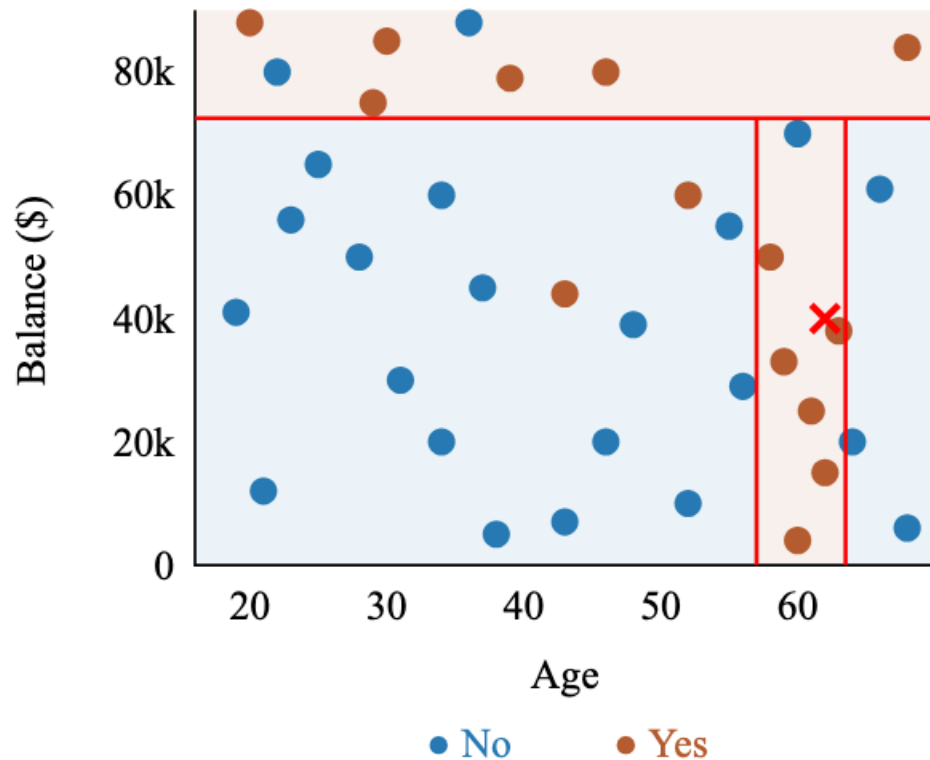
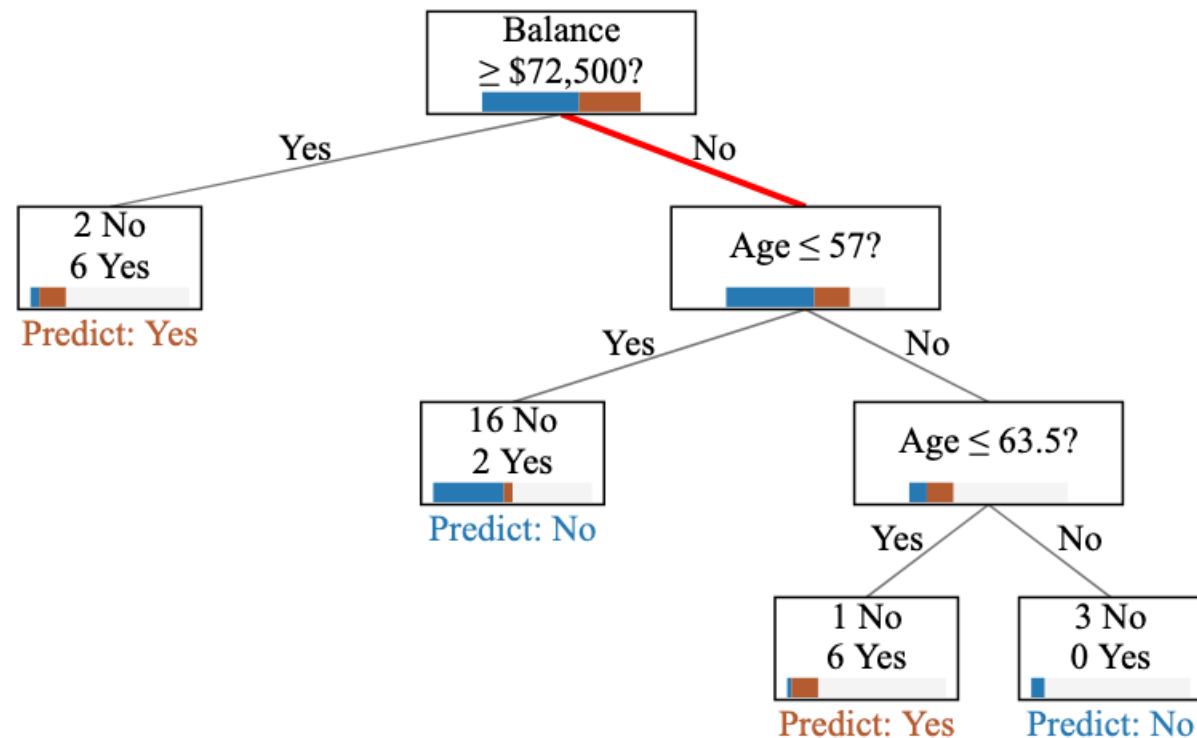
fit: Learn the Questions



predict: Follow the Questions

Training Builds the Rules; Prediction Uses Them

Age 62 · Balance \$40,000

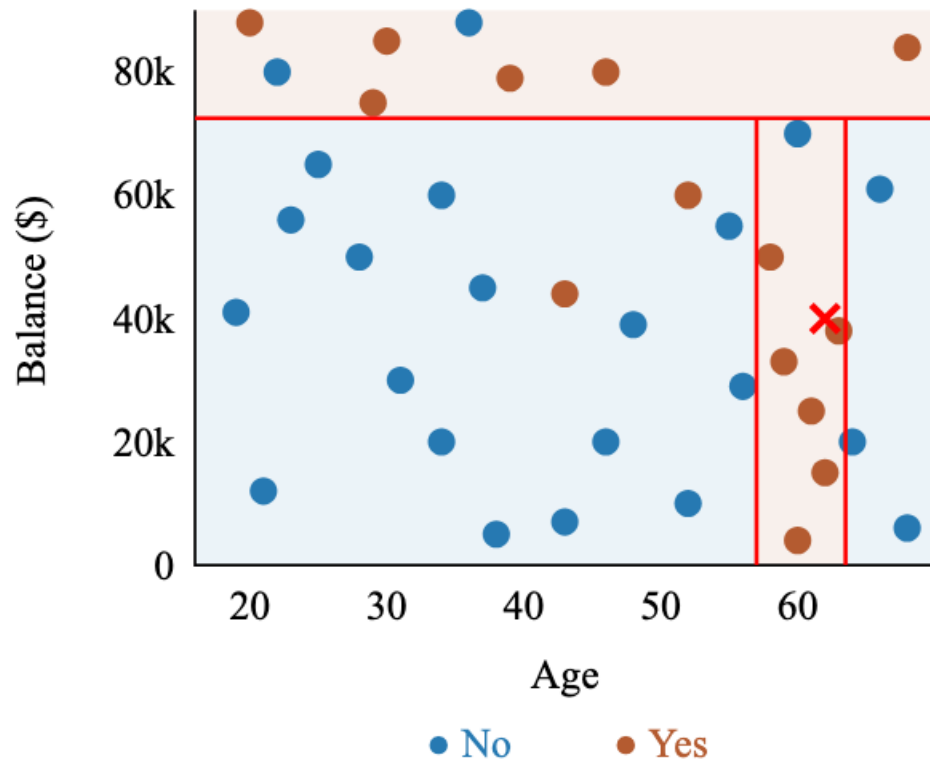
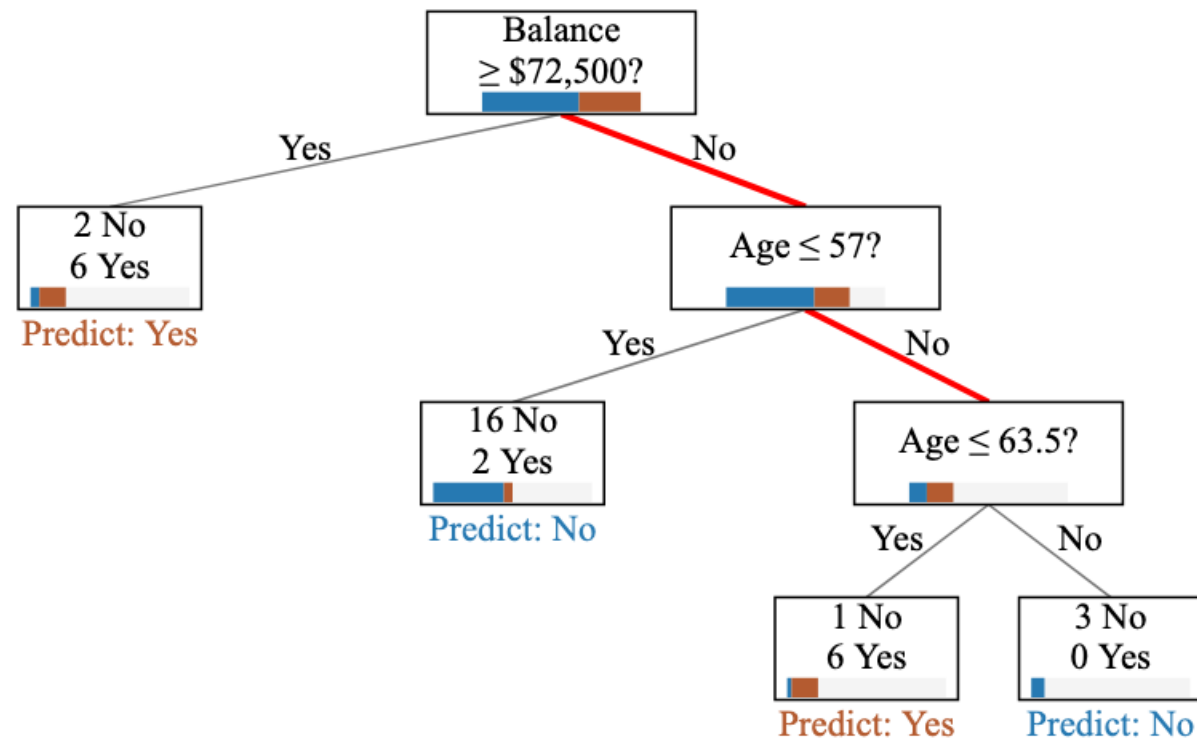


fit: Learn the Questions

predict: Follow the Questions

Training Builds the Rules; Prediction Uses Them

Age 62 · Balance \$40,000

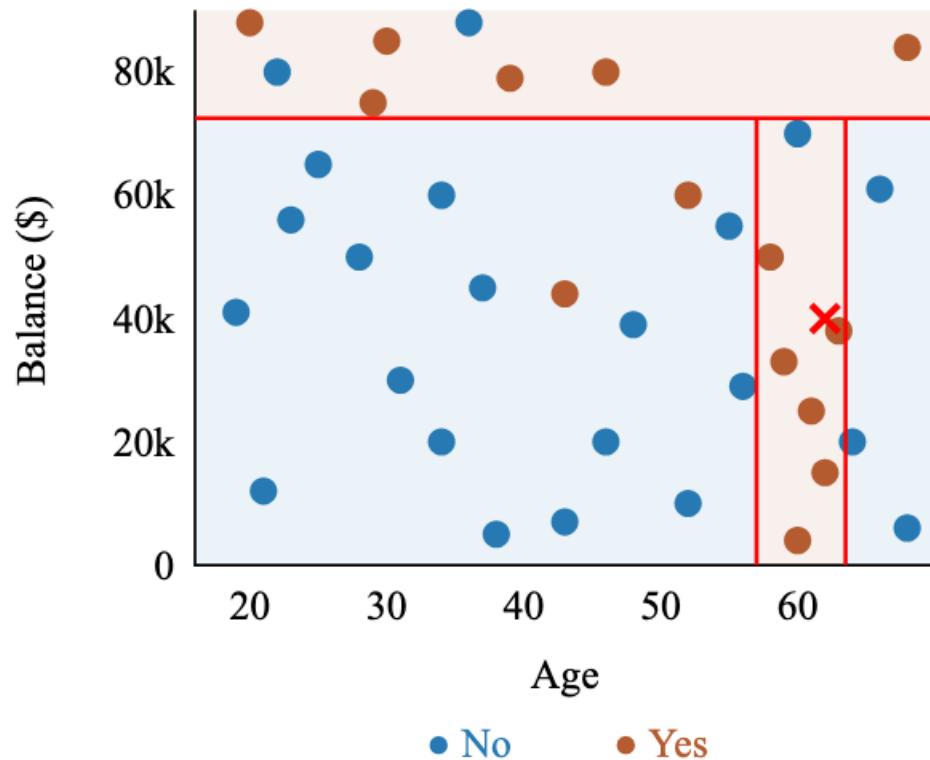


fit: Learn the Questions

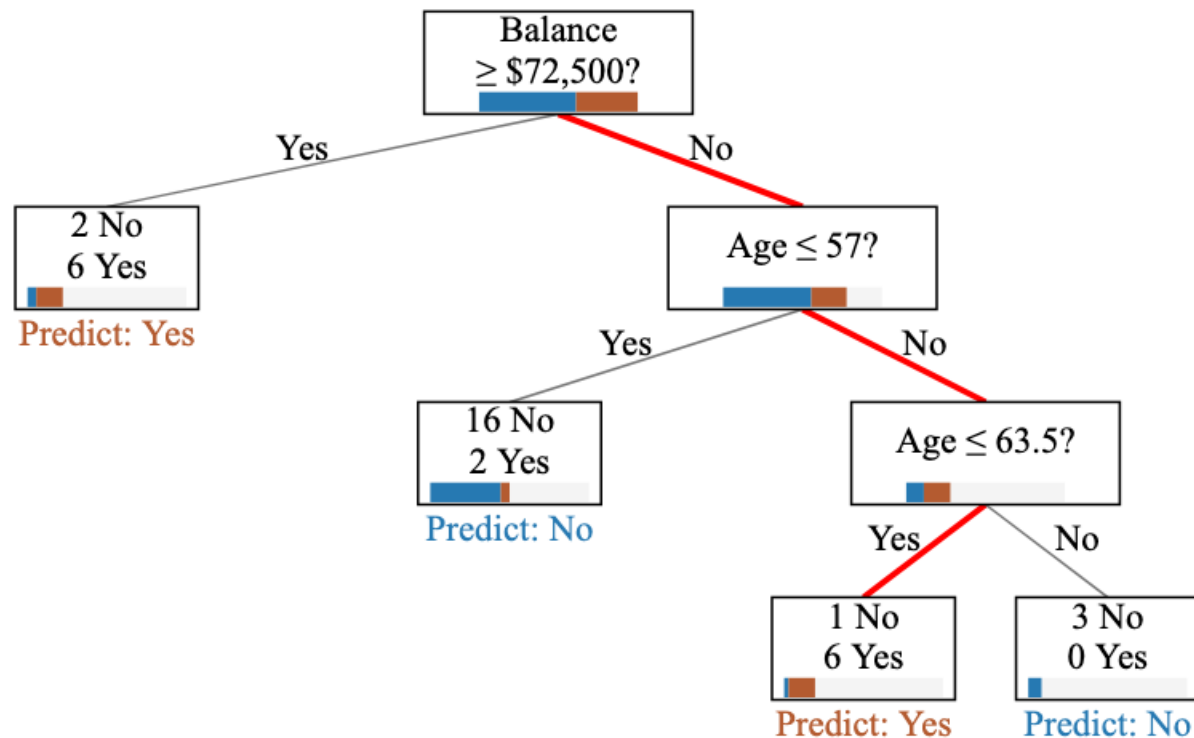
predict: Follow the Questions

Training Builds the Rules; Prediction Uses Them

Age 62 · Balance \$40,000



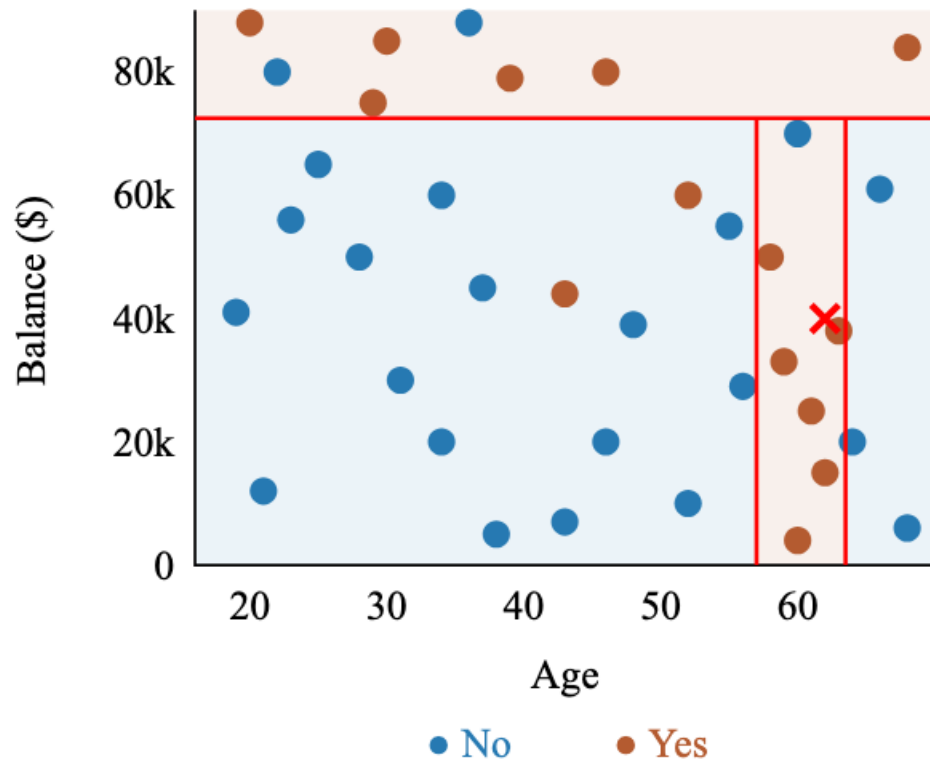
fit: Learn the Questions



predict: Follow the Questions

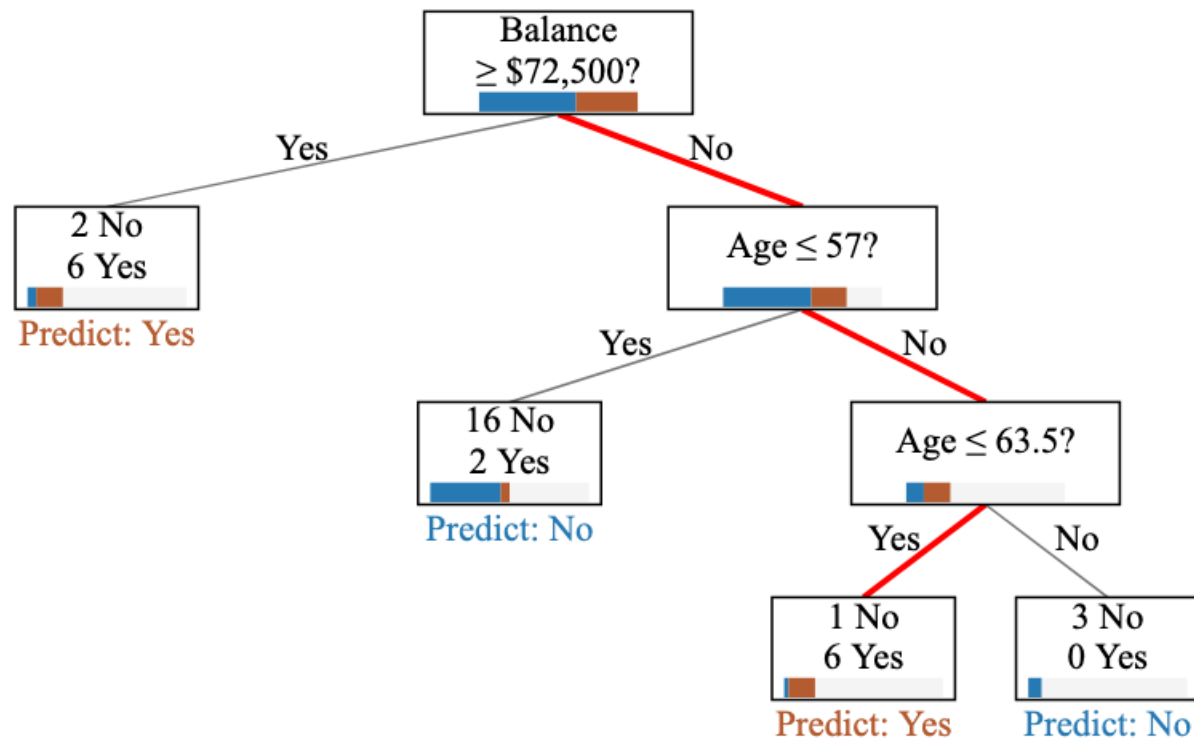
Training Builds the Rules; Prediction Uses Them

Age 62 · Balance \$40,000



Predict: Yes

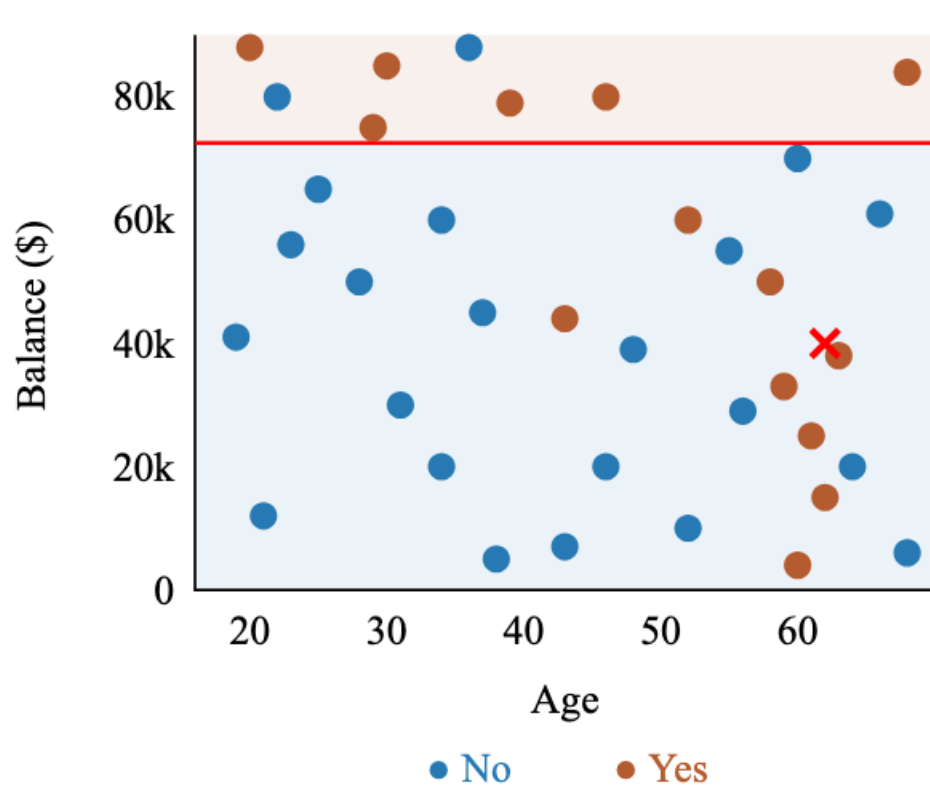
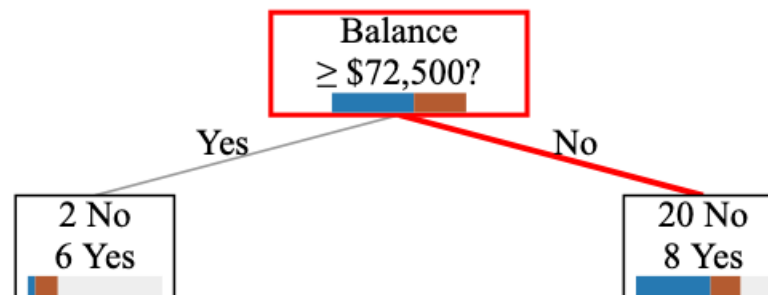
fit: Learn the Questions



predict: Follow the Questions

How Many Questions Should the Tree Ask?

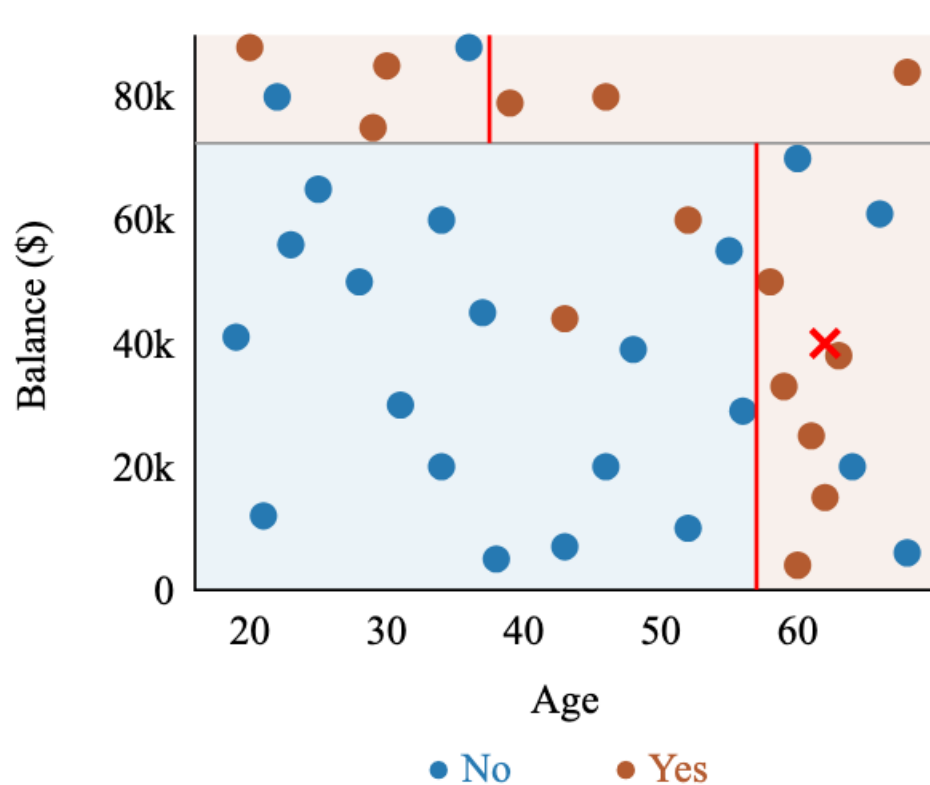
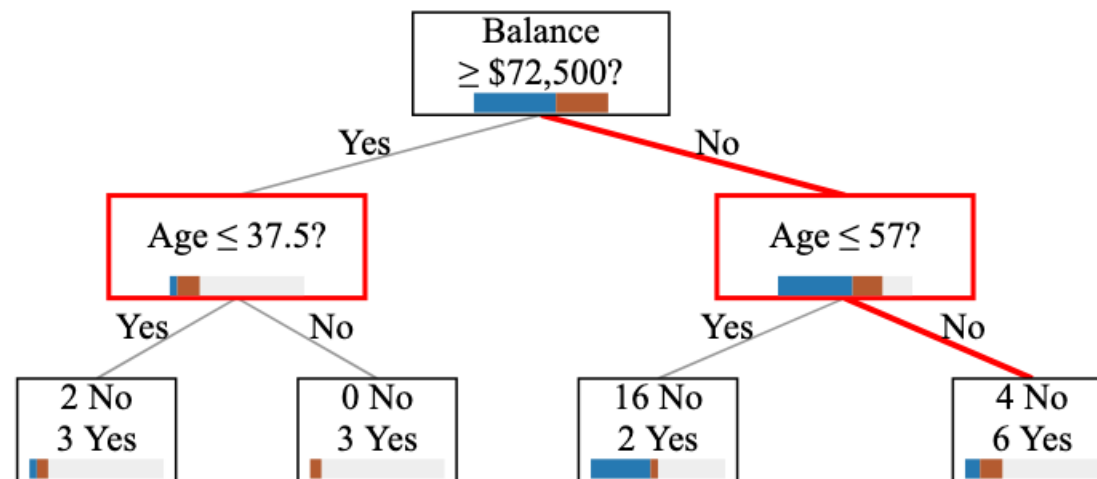
Maximum Depth 1 · At Most 1 Question Along a Path



Predict: No

How Many Questions Should the Tree Ask?

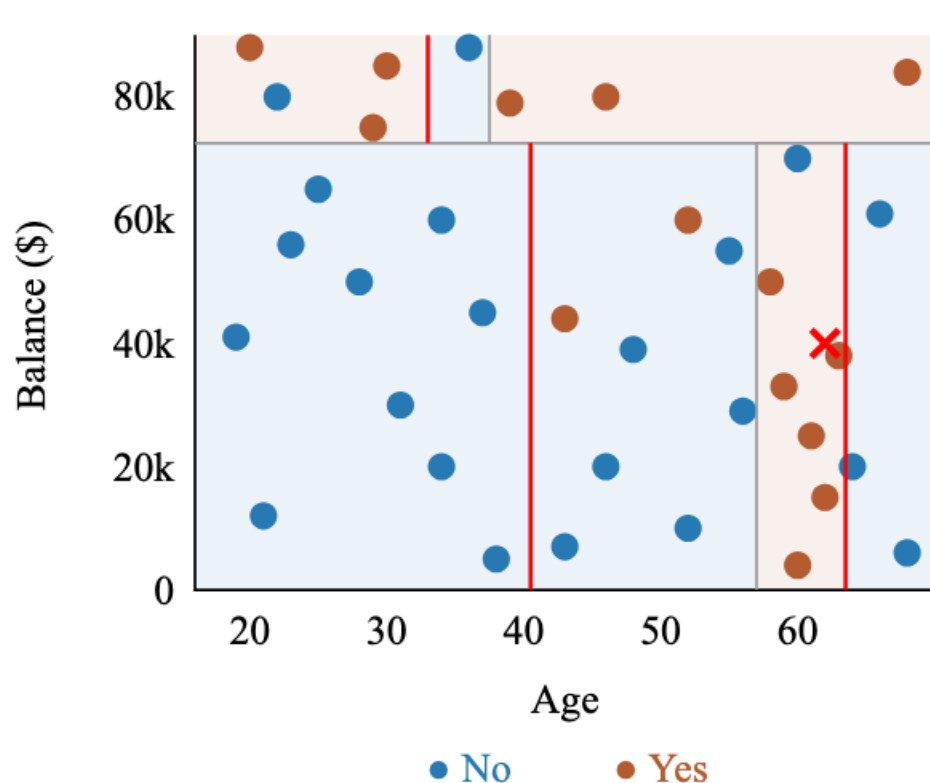
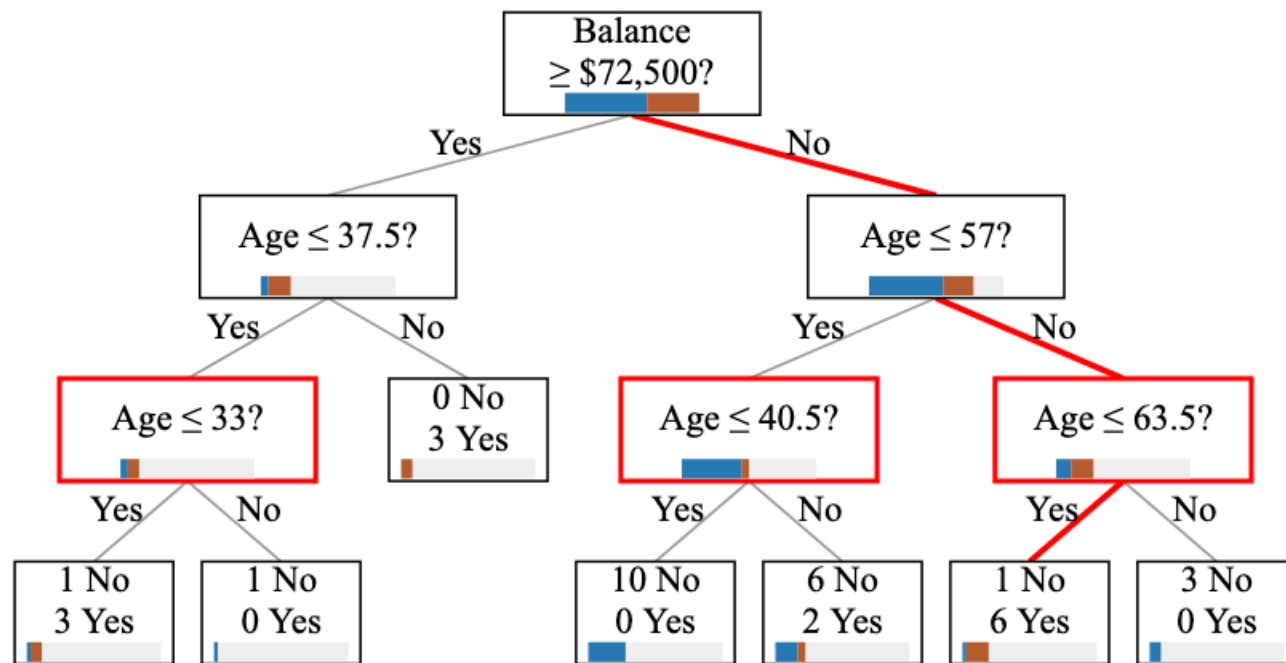
Maximum Depth 2 · At Most 2 Questions Along a Path



Predict: Yes

How Many Questions Should the Tree Ask?

Maximum Depth 3 · At Most 3 Questions Along a Path



Predict: Yes

More Questions Allow More Detailed Rules

maib7002.dirk.hk/L4.zip

Demo: `decision_tree_sklearn.ipynb`

Activity: Follow a Tree Path

Notebook: [l4/decision_tree_sklearn.ipynb](#)

1. Choose one leaf and explain every question on its path
2. Predict what changes when maximum depth goes from 4 to 2
3. Run the change and explain the new regions

Submit the brief check-in on Ed

Why Combine Different Trees?

1. One tree may be misled by unusual training examples

Why Combine Different Trees?

1. One tree may be misled by unusual training examples
2. Different trees can make different mistakes

Why Combine Different Trees?

1. One tree may be misled by unusual training examples
2. Different trees can make different mistakes
3. Combining them can reduce the influence of any one tree's mistakes

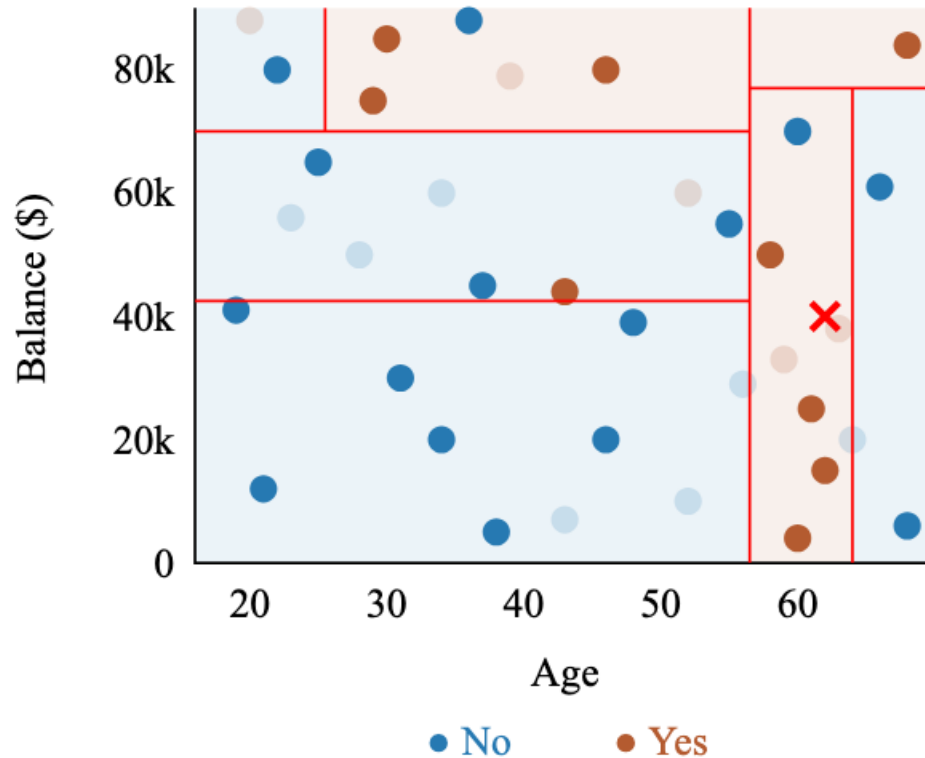
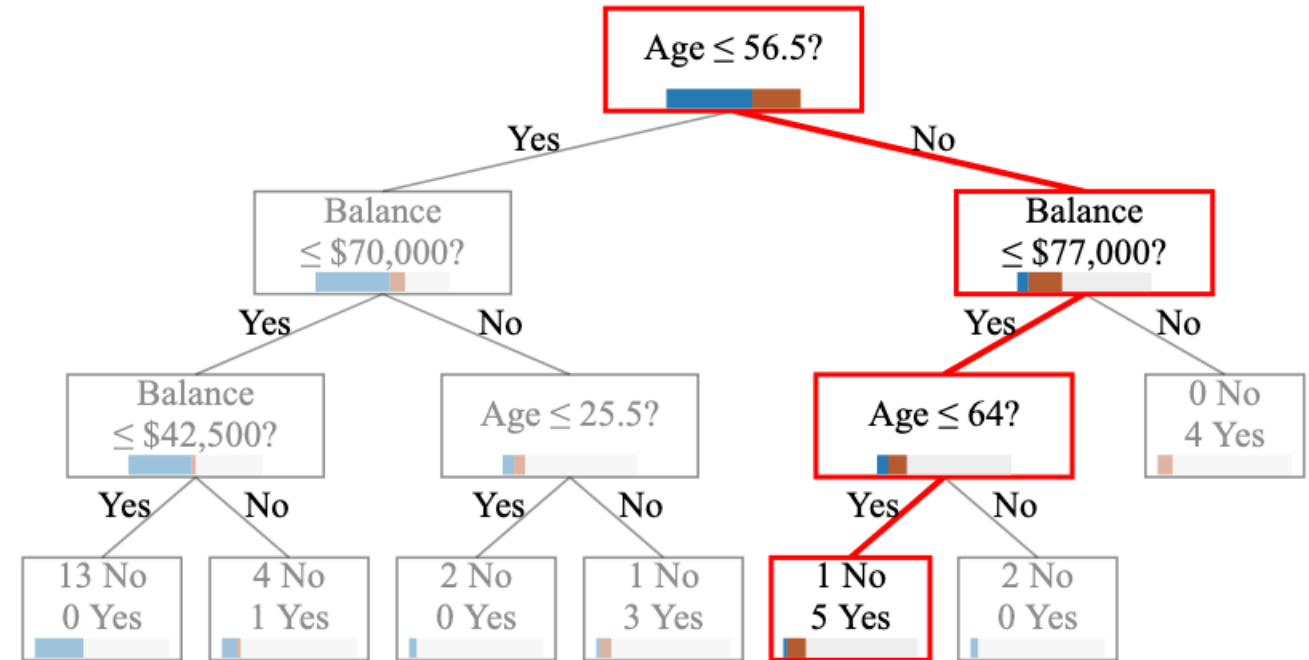
Why Combine Different Trees?

1. One tree may be misled by unusual training examples
2. Different trees can make different mistakes
3. Combining them can reduce the influence of any one tree's mistakes

Our Goal: Better Predictions for New Customers

Why Rely on Just One Tree?

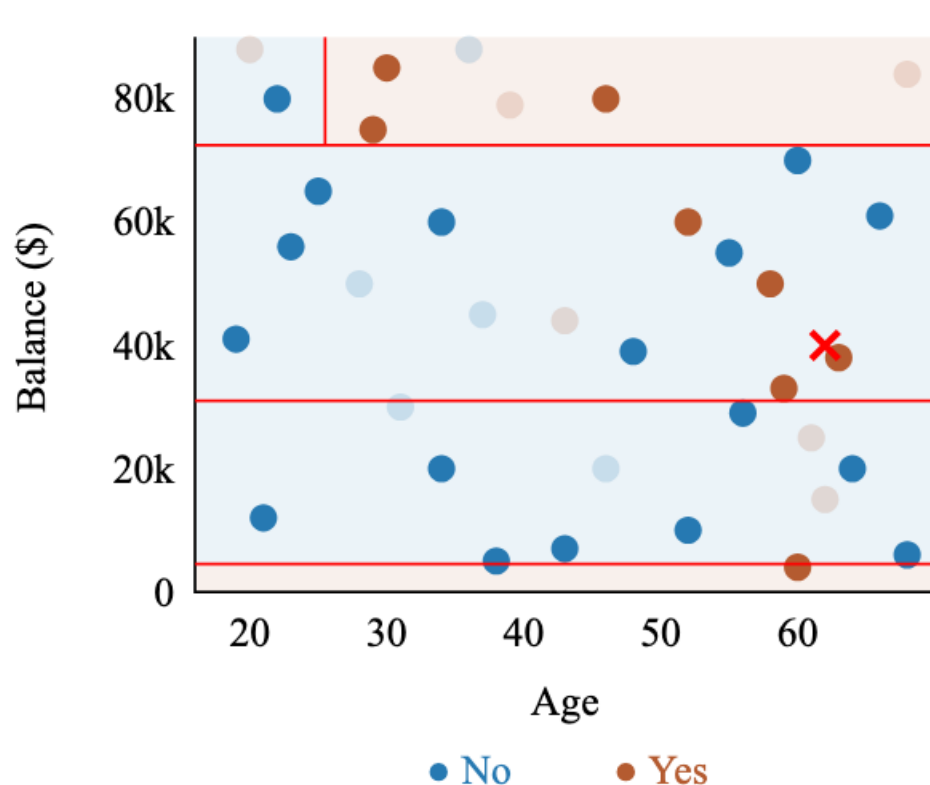
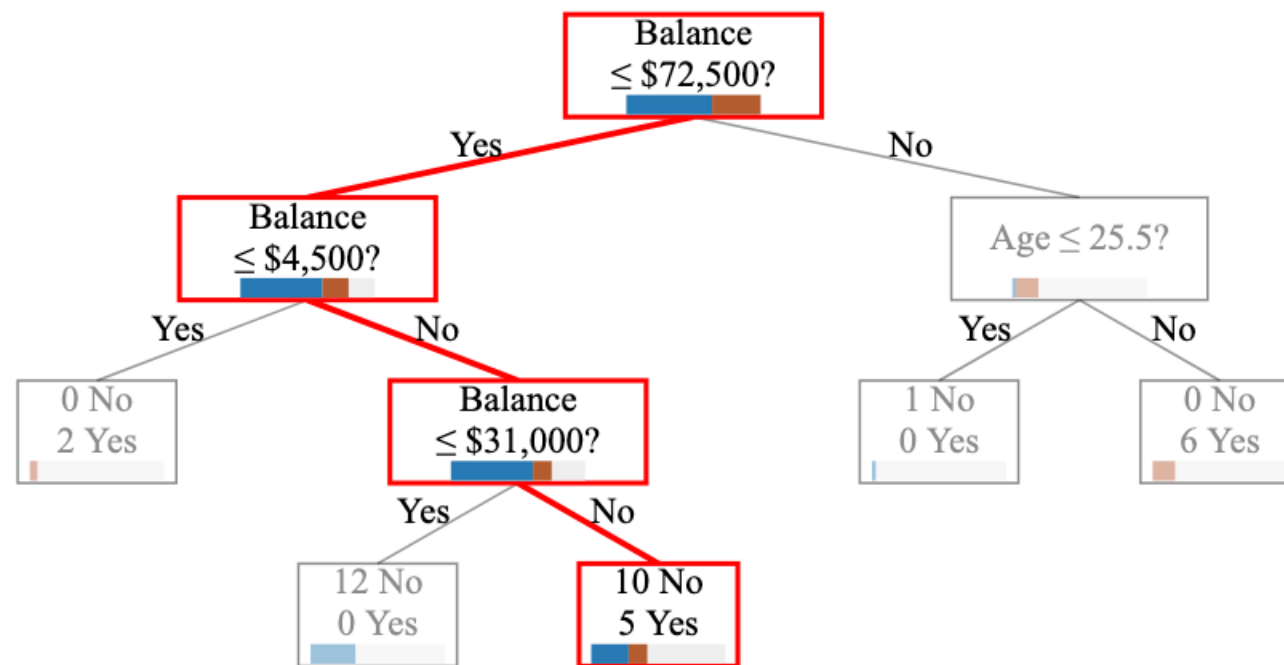
Tree 1: Draw Customers From Our Dataset and Train



Age 62 · \$40,000 → Predict: Yes

Why Rely on Just One Tree?

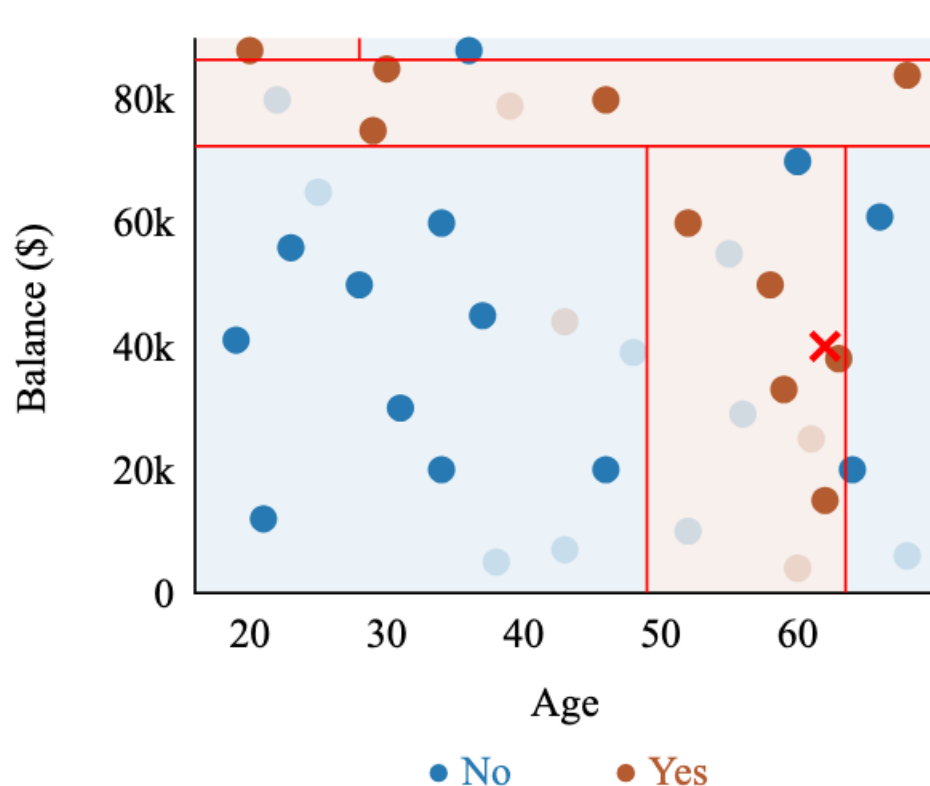
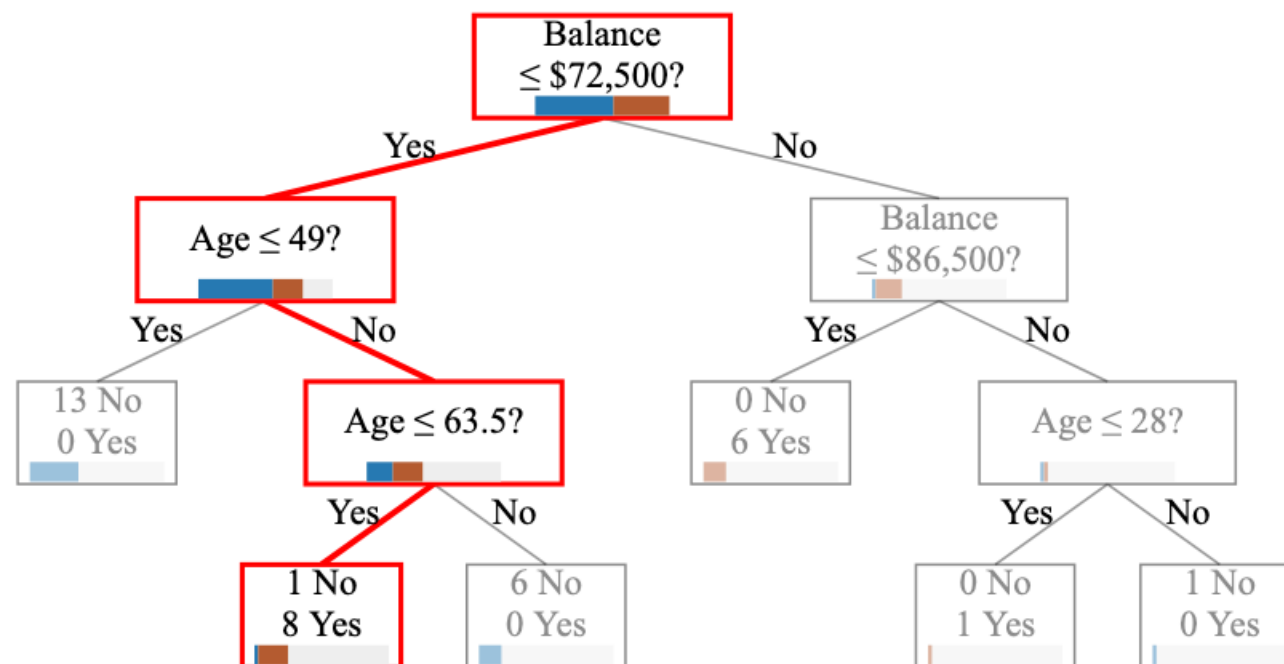
Tree 2: Draw Again and Train a New Tree



Age 62 · \$40,000 → Predict: No

Why Rely on Just One Tree?

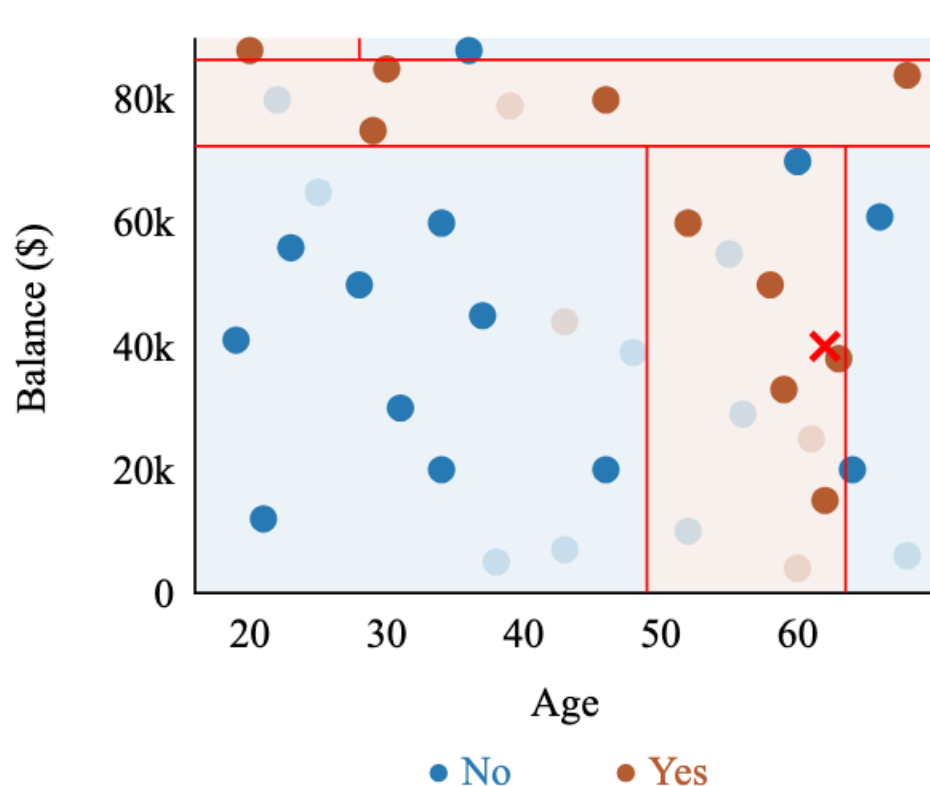
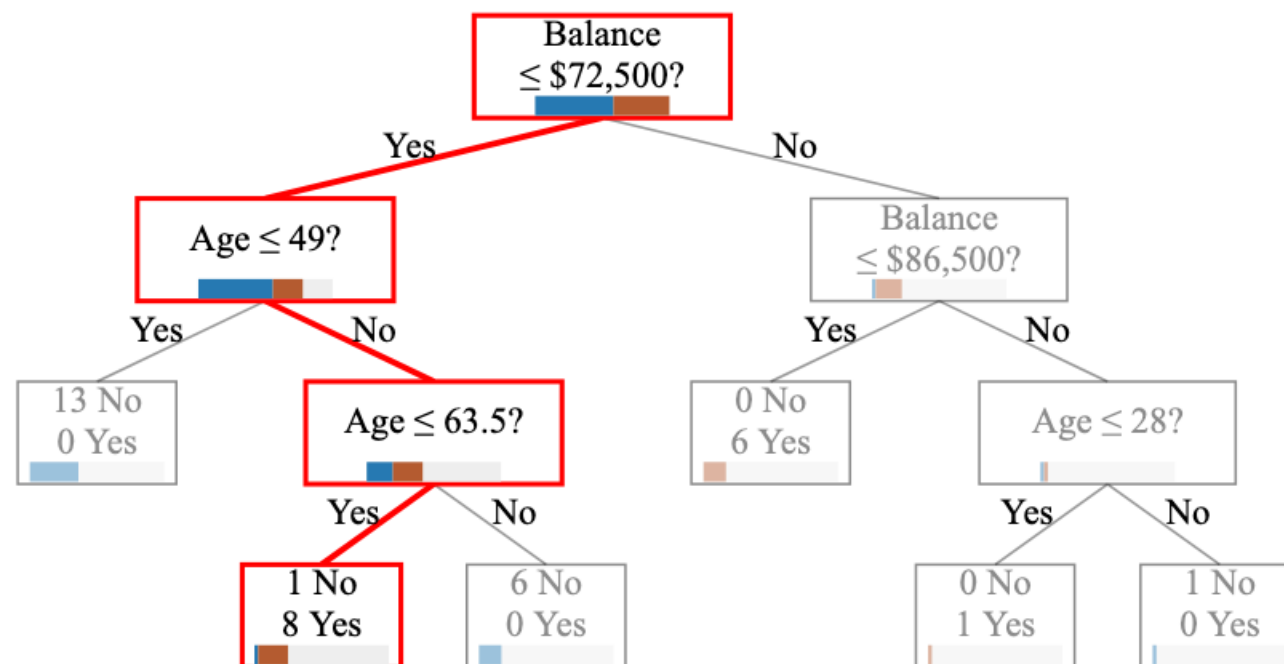
Tree 3: Repeat With Another Random Draw



Age 62 · \$40,000 → Predict: Yes

Why Rely on Just One Tree?

Tree 3: Repeat With Another Random Draw

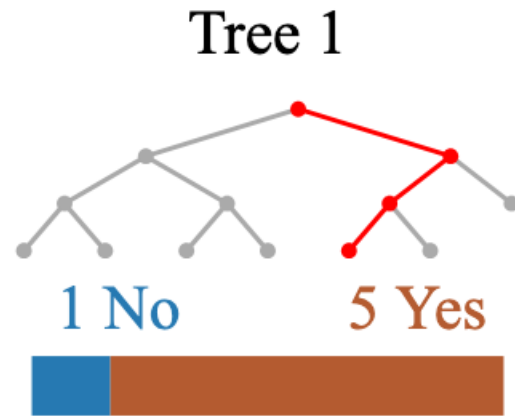
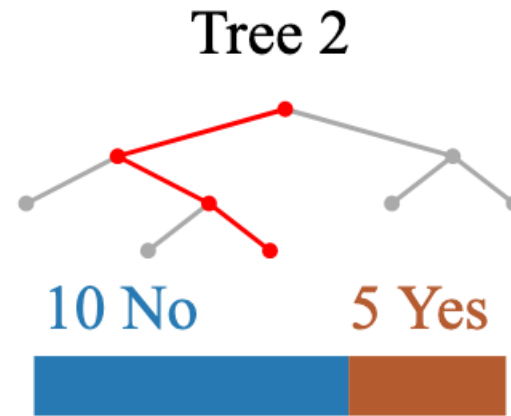
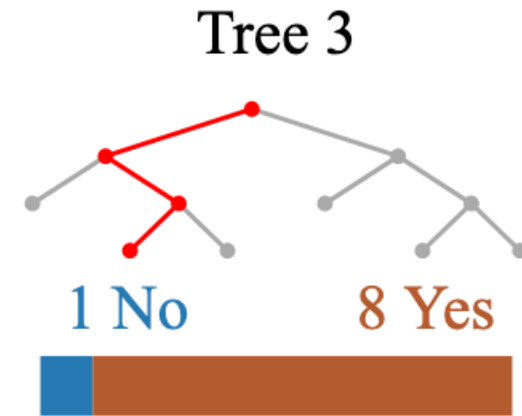


Age 62 · \$40,000 → Predict: Yes

Can Combining Different Trees Make the Prediction More Stable?

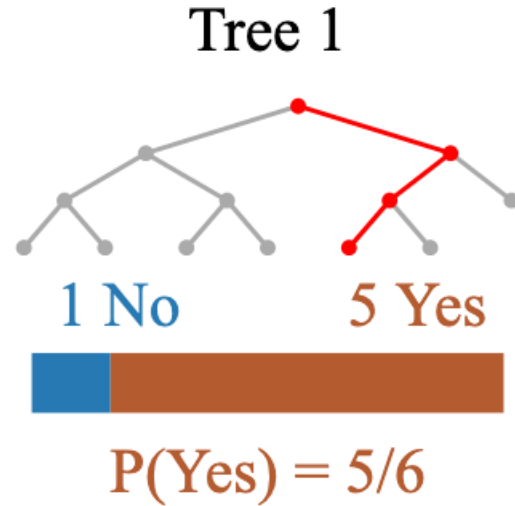
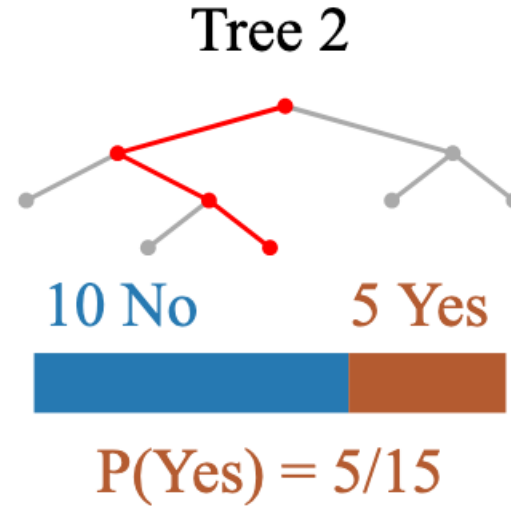
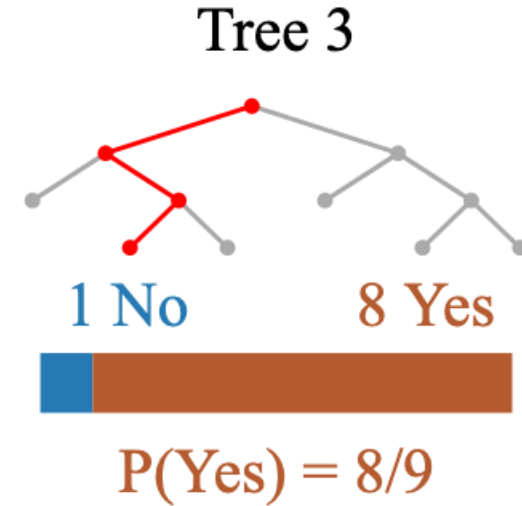
Combine the Three Trees

Same Customer: Age 62 · Balance \$40,000



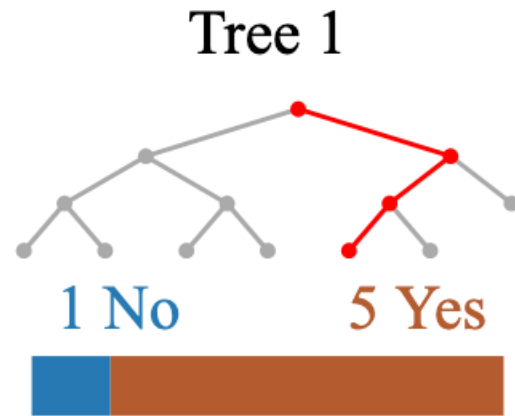
Combine the Three Trees

Same Customer: Age 62 · Balance \$40,000

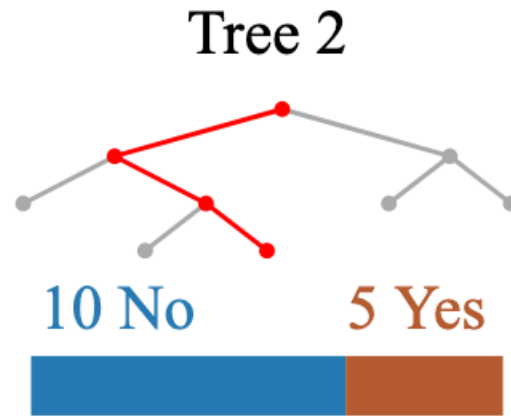


Combine the Three Trees

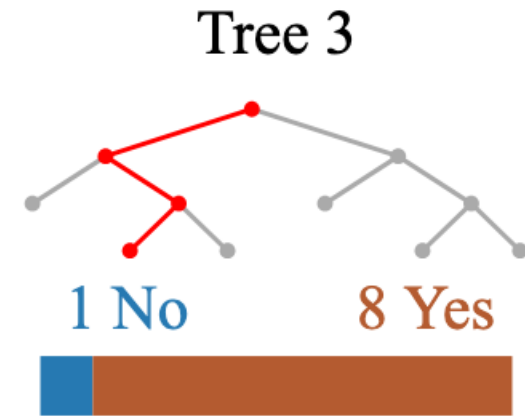
Same Customer: Age 62 · Balance \$40,000



$$P(\text{Yes}) = 5/6$$



$$P(\text{Yes}) = 5/15$$



$$P(\text{Yes}) = 8/9$$

$$\text{Average } P(\text{Yes}) = (5/6 + 5/15 + 8/9) \div 3 \approx 68.5\%$$



Combined Prediction: Yes

How Can We Make the Trees Different?

Make the trees different in two ways:

1. Use different training samples

How Can We Make the Trees Different?

Make the trees different in two ways:

1. Use different training samples
2. Consider different features at each split

Random Forest Settings in sklearn

Setting

`max_samples=20`

Random Choice and Effect

Randomly draw 20 records per tree; repeats are allowed
Fewer draws → each tree sees less training data

Random Forest Settings in sklearn

Setting

Random Choice and Effect

`max_samples=20`

Randomly draw 20 records per tree; repeats are allowed
Fewer draws → each tree sees less training data

`max_features=2`

Randomly select 2 features; split using the best 1
Fewer candidates → a smaller random subset
Using all features removes this source of randomness

Random Forest Settings in sklearn

Setting	Random Choice and Effect
<code>max_samples=20</code>	Randomly draw 20 records per tree; repeats are allowed Fewer draws → each tree sees less training data
<code>max_features=2</code>	Randomly select 2 features; split using the best 1 Fewer candidates → a smaller random subset Using all features removes this source of randomness
<code>n_estimators=25</code>	Build 25 trees, each with its own random draws More trees → average across more random draws

Random Forest Settings in sklearn

Setting

Random Choice and Effect

`max_samples=20`

Randomly draw 20 records per tree; repeats are allowed
Fewer draws → each tree sees less training data

`max_features=2`

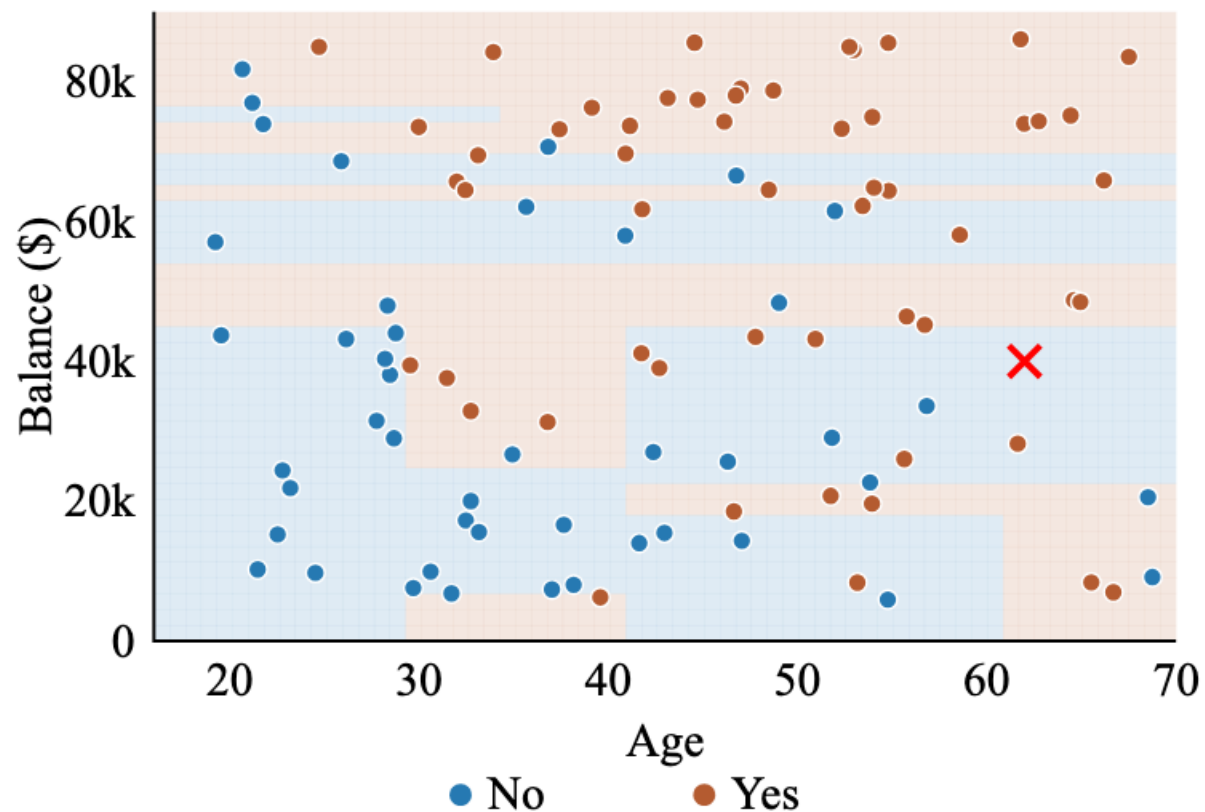
Randomly select 2 features; split using the best 1
Fewer candidates → a smaller random subset
Using all features removes this source of randomness

`n_estimators=25`

Build 25 trees, each with its own random draws
More trees → average across more random draws

`random_state=1` makes the random draws repeatable

Do More Trees Make Predictions More Stable?



Number of Trees

New Simulated Customers · Deep Trees

☒ 1 ☐ 3 ☐ 5 ☐ 10 ☐ 50

Redraw the Trees

Draw 1



Age 62 · \$40,000

$P(\text{No}) = 100.0\%$

$P(\text{Yes}) = 0.0\%$

Predict: No

maib7002.dirk.hk/L4.zip

Demo: random_forests_sklearn.ipynb

Group Activity: Random Forests

Notebook: [l4/random_forests_sklearn.ipynb](#)

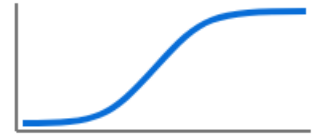
1. Run the notebook with 25 trees
2. Change to 1 tree; rerun the fit, plot and prediction cells
3. Compare the probabilities and predicted species

Discuss and Submit One Group Check-In on Ed

Supervised Learning: L3 and L4

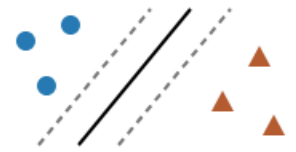
Logistic Regression

Estimate Class Probabilities



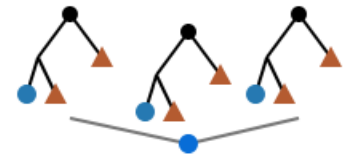
SVM

Separate Classes with a Margin



Random Forest

Learn Questions, Then Combine Trees



kNN

Use the Labels of Nearby Samples



kNN Keeps Examples and Looks for Neighbours

1. Keep the labelled customers

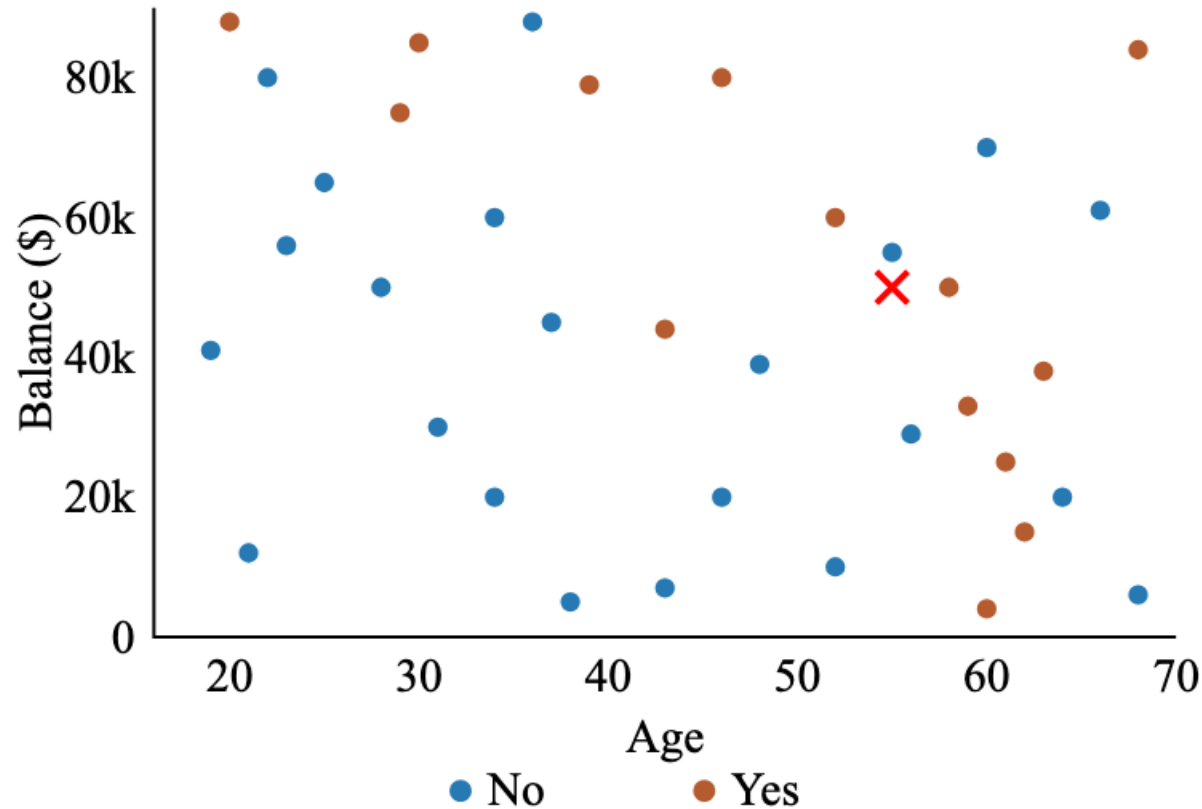
kNN Keeps Examples and Looks for Neighbours

1. Keep the labelled customers
2. Find neighbours when a new customer needs a prediction

kNN Keeps Examples and Looks for Neighbours

1. Keep the labelled customers
2. Find neighbours when a new customer needs a prediction
3. Use comparable feature scales to measure distance

What Did Similar Customers Do?



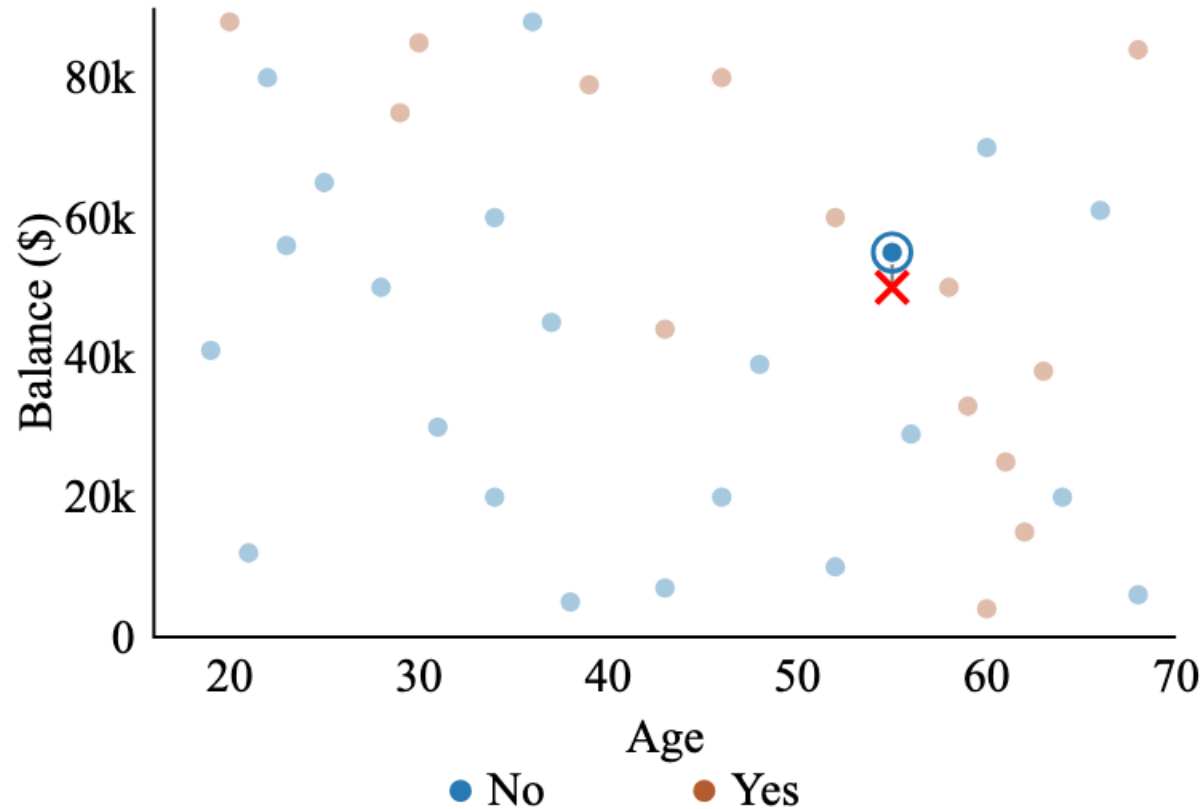
Age 55 · \$50,000

Will This Customer Say Yes?

kNN = k-Nearest Neighbours

Distance Uses Scaled Age and Balance

What Did Similar Customers Do?



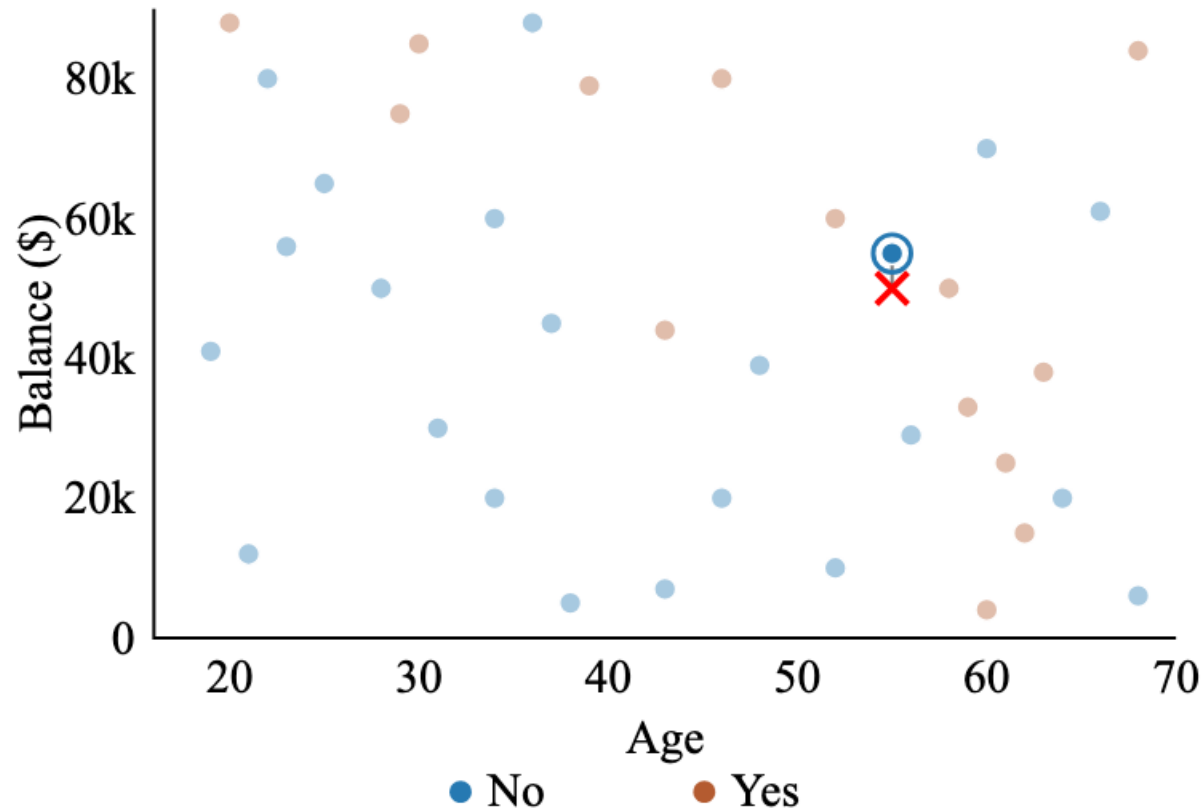
Age 55 · \$50,000

k = 1: One Neighbour

What Is the Nearest Customer's Label?

Distance Uses Scaled Age and Balance

What Did Similar Customers Do?



Age 55 · \$50,000

$k = 1$: One Neighbour

No

Yes

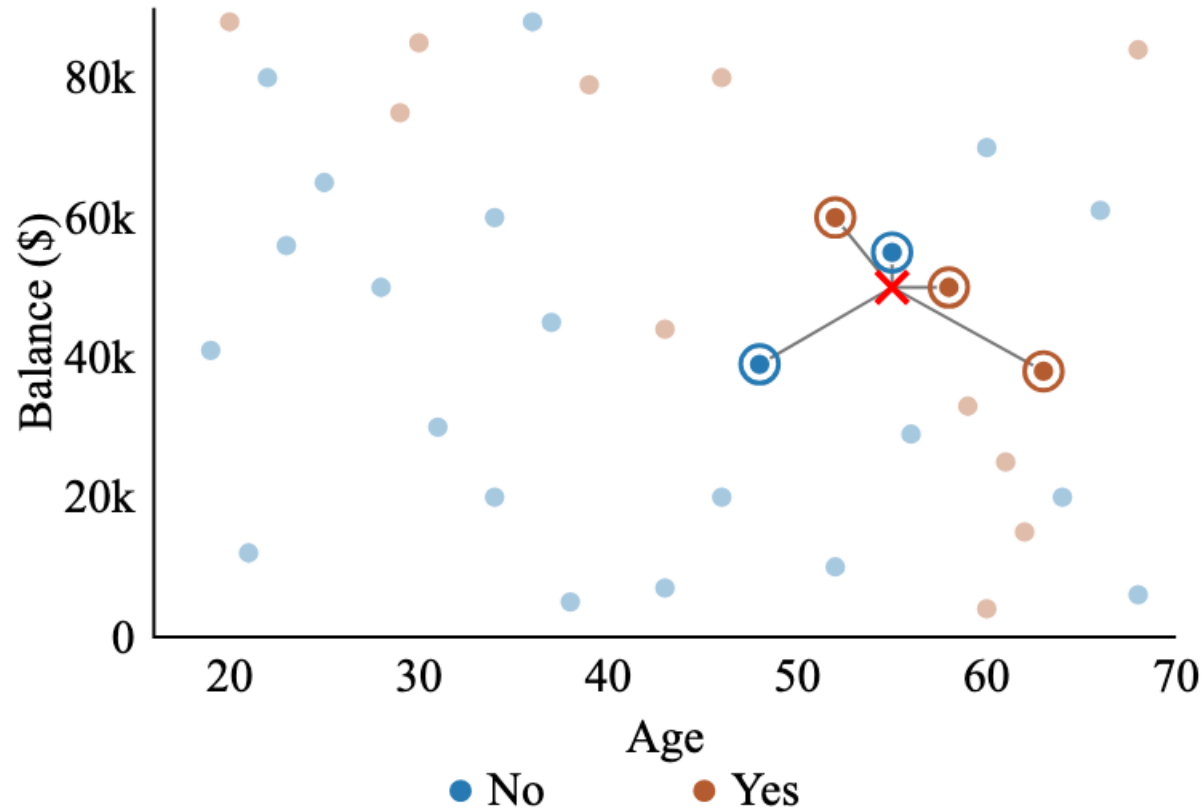


Predict: No

One Customer, One Vote

Distance Uses Scaled Age and Balance

What Did Similar Customers Do?



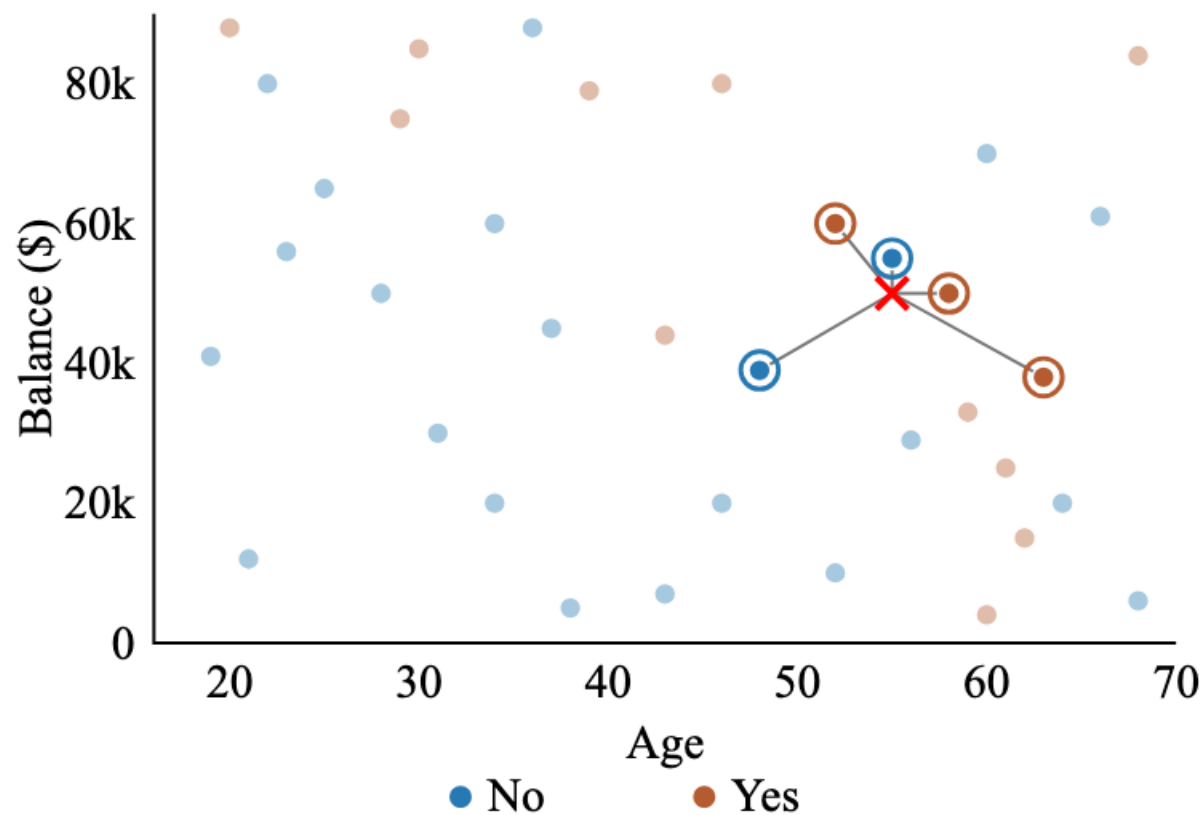
Age 55 · \$50,000

k = 5: Five Neighbours

Count the No and Yes Labels

Distance Uses Scaled Age and Balance

What Did Similar Customers Do?



Age 55 · \$50,000

k = 5: Five Neighbours

No

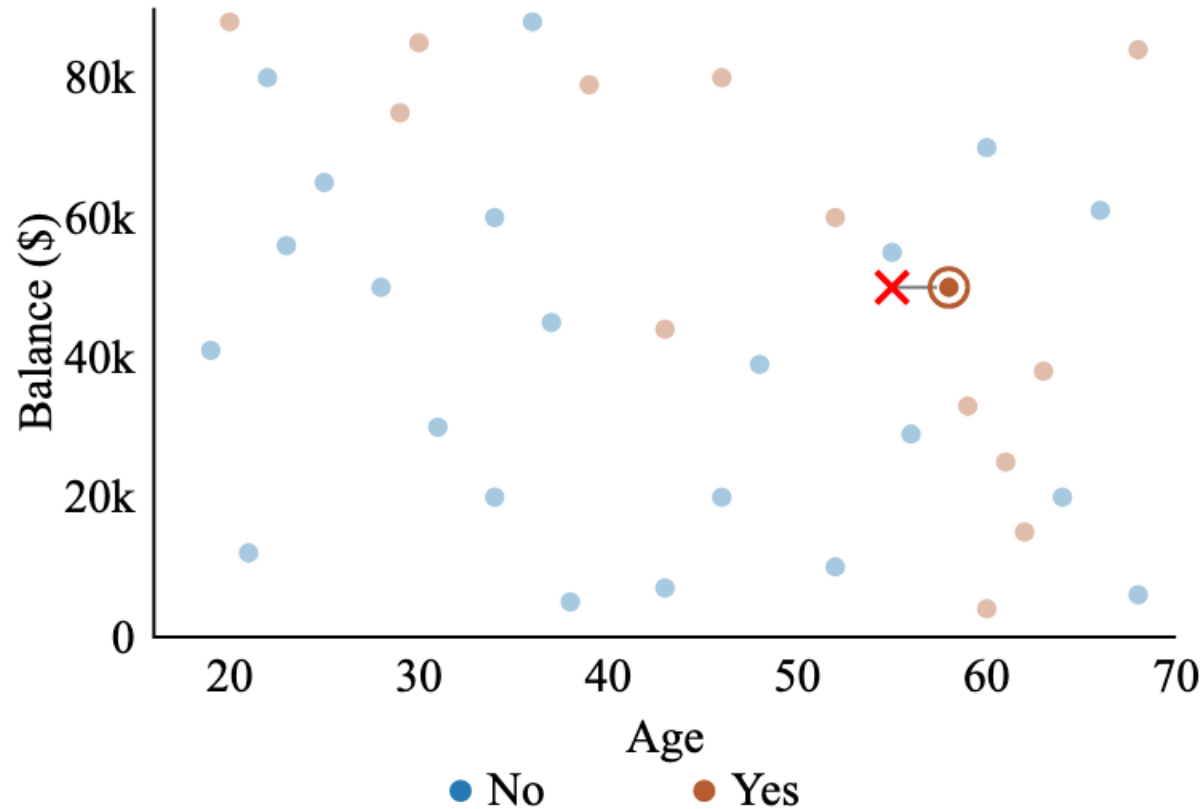


Yes

Predict: **Yes**

The Most Common Label Wins

What Makes a Customer Near?



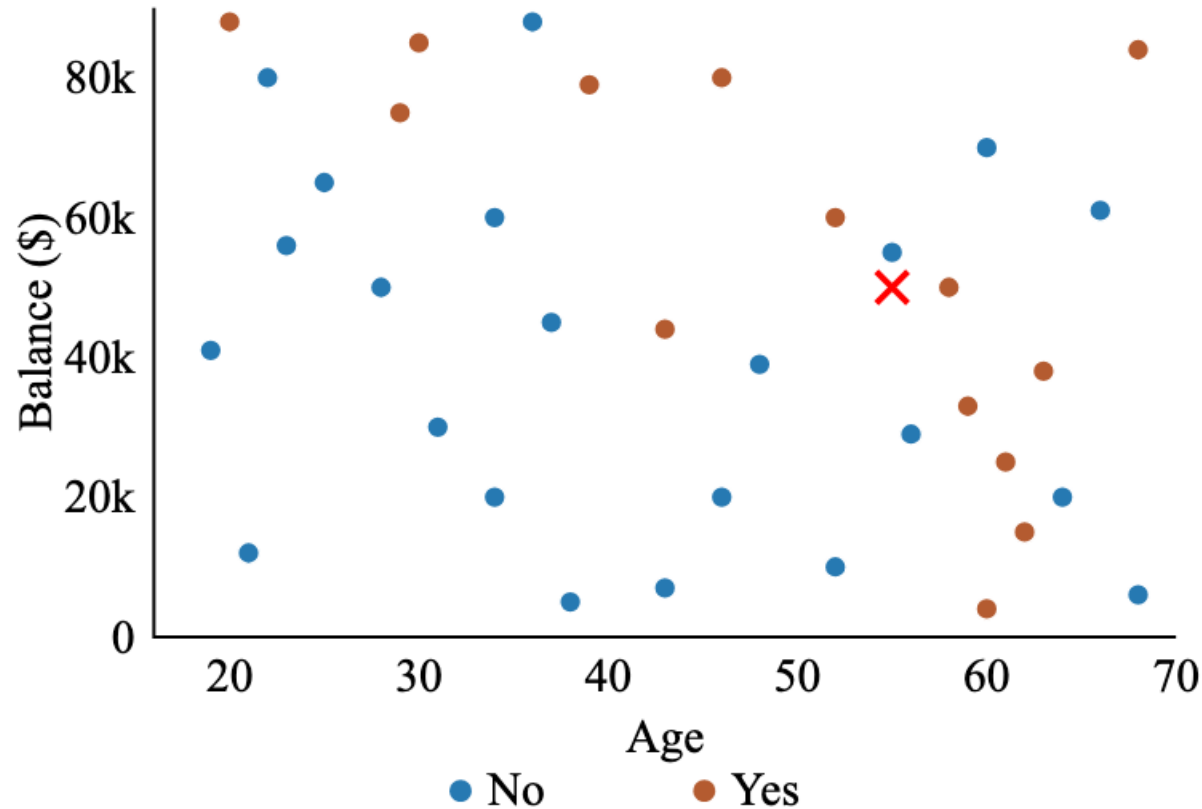
Age 55 · \$50,000

Without Scaling

Nearest: Age 58 · \$50,000

Dollar Differences Dominate Year Differences

What Makes a Customer Near?



Age 55 · \$50,000

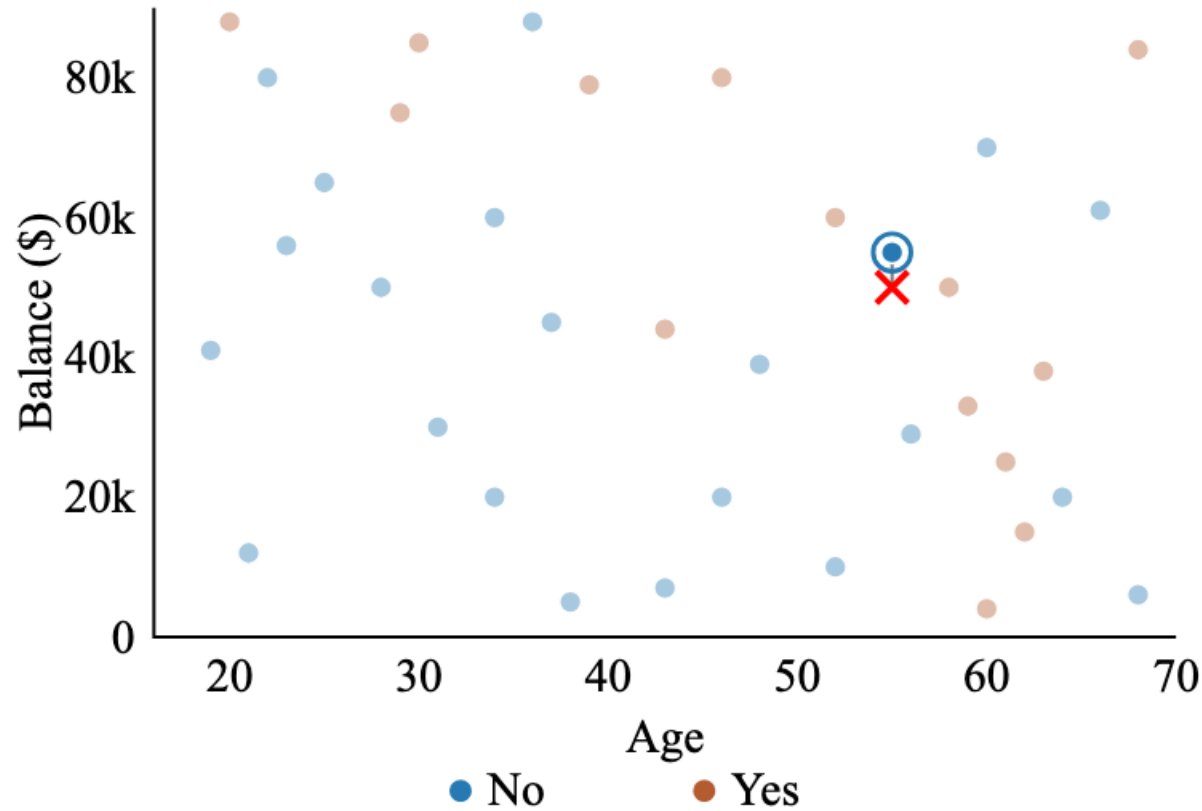
Scale Age and Balance

Divide by Each Feature's Typical Spread

Age Difference $\div$ 15.5 Years

Balance Difference $\div$ \$26,764

What Makes a Customer Near?



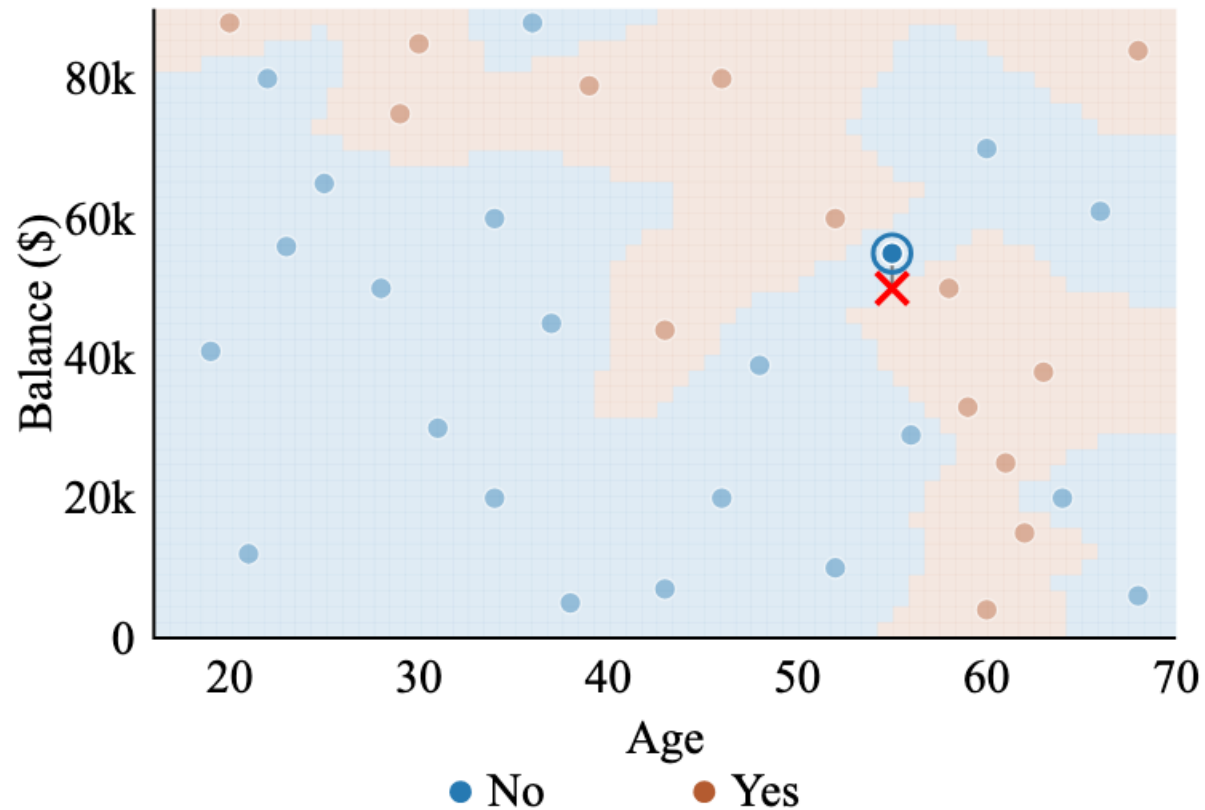
Age 55 · \$50,000

After Scaling

Nearest: Age 55 · \$55,000

Same Customers, a Different Distance Rule

How Many Neighbours Should Vote?



Age 55 · \$50,000

$k = 1$ Neighbour

No

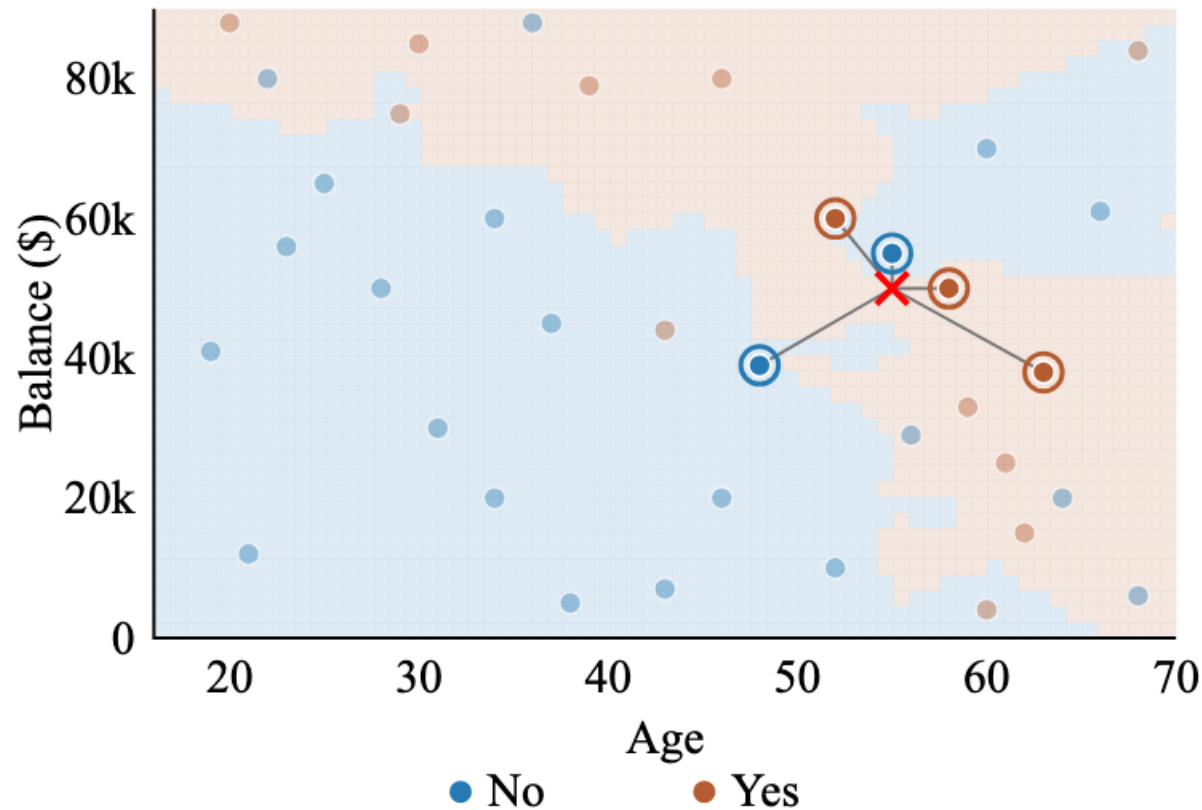
Yes



Predict: No

One Label Can Create a Small Region

How Many Neighbours Should Vote?



Age 55 · \$50,000

k = 5 Neighbours

No

Yes



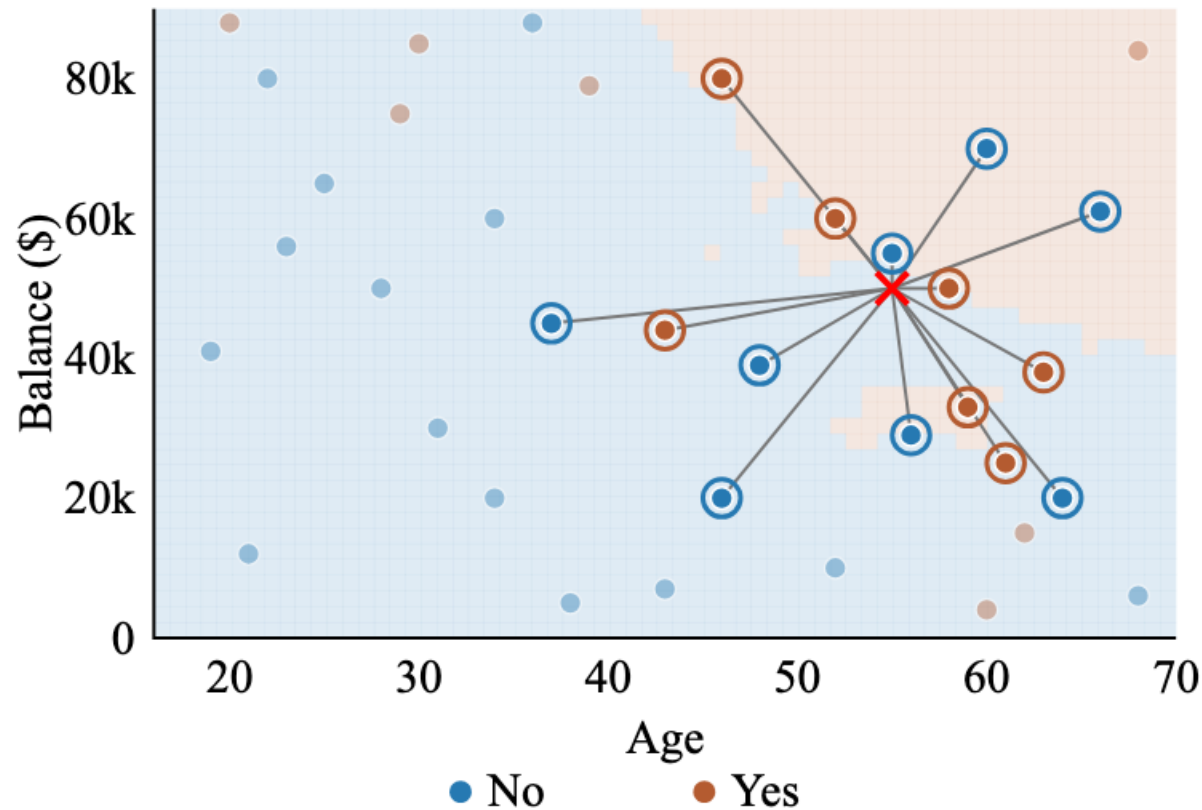
0

5 Votes

Predict: Yes

More Neighbours Share the Vote

How Many Neighbours Should Vote?



Age 55 · \$50,000

k = 15 Neighbours

No

Yes



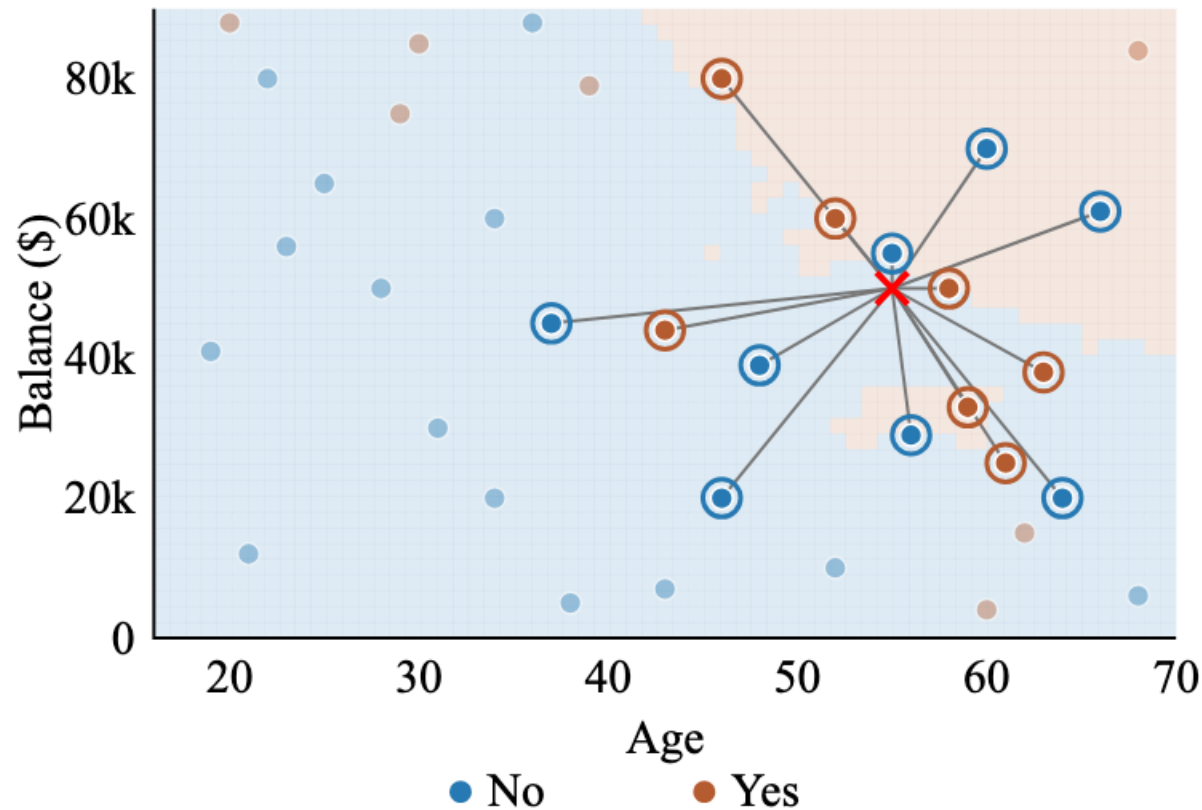
0

15 Votes

Predict: No

A Wider Neighbourhood Can Lose Local Detail

How Many Neighbours Should Vote?



Age 55 · \$50,000

k = 15 Neighbours

k = 1 k = 3 k = 5 k = 7 **k = 15**

No

Yes



0

15 Votes

Predict: No

Move × to Explore Another Customer

The Main kNN Settings

Setting

Meaning in Our Example

`n_neighbors=5`

`n_neighbors` is k , the number of neighbours
 $k = 5$: Ask the five nearest customers

The Main kNN Settings

Setting	Meaning in Our Example
<code>n_neighbors=5</code>	<code>n_neighbors</code> is k , the number of neighbours $k = 5$: Ask the five nearest customers
<code>weights="uniform"</code>	Give each neighbour one equal vote

The Main kNN Settings

Setting	Meaning in Our Example
<code>n_neighbors=5</code>	<code>n_neighbors</code> is k , the number of neighbours $k = 5$: Ask the five nearest customers
<code>weights="uniform"</code>	Give each neighbour one equal vote
<code>metric="euclidean"</code>	Measure straight-line distance after scaling

The Main kNN Settings

Setting	Meaning in Our Example
<code>n_neighbors=5</code>	<code>n_neighbors</code> is k , the number of neighbours $k = 5$: Ask the five nearest customers
<code>weights="uniform"</code>	Give each neighbour one equal vote
<code>metric="euclidean"</code>	Measure straight-line distance after scaling

For the Activity, Change Only `n_neighbors`

maib7002.dirk.hk/L4.zip

Demo: knn_sklearn.ipynb

Group Activity: kNN

Notebook: [l4/knn_sklearn.ipynb](#)

1. Run the notebook with `n_neighbors=5`
2. Predict the answer with one neighbour, then change to `1`
3. Rerun Step 3 and Step 4, then explain the vote

Discuss and Submit One Group Check-In on Ed

More Supervised Models in sklearn

Model Family	Examples	Another Way to Predict
Naive Bayes	<code>GaussianNB</code> , <code>MultinomialNB</code>	Combine Feature Evidence Using Probabilities
Boosting	<code>HistGradientBoostingClassifier</code>	Add Models That Improve the Current Fit
Neural Networks	<code>MLPClassifier</code>	Learn Layers of Nonlinear Transformations
Regression Models	<code>Ridge</code> , <code>RandomForestRegressor</code>	Predict a Number Rather Than a Class

Recognize the Options; Understand the Models You Use

Group Project: Explain and Critique

Find a Partner Group · Suggested Time: 30 Minutes

1. Explain Your Current Project (5 min)

User and decision, example input and output, data source, baseline and evaluation

2. Listen and Critique (5 min)

Give one strength, one concern, and one suggested improvement

3. Swap Roles (10 min)

Repeat the explanation and feedback

4. Refine Your Plan (7 min)

Agree on one concrete improvement within your own group

Last Week's Proposal Submissions

Status	Groups
Submitted (26)	B, C, D, E, F, G, I, J, K, L, M, O, P, Q, R, S, T, W, Y, Z, Alpha, Gamma, Delta, Epsilon, Zeta, Eta
No Submission (6)	A, H, N, U, V, X

Submit Your
Exit Ticket on Ed

Exit Ticket (3 min): Your Group, Partner Group, Feedback Received, and Planned Improvement

Next Time

Evaluation and Unsupervised Learning

*Machine Learning with PyTorch and
Scikit-Learn*

Raschka, Liu and Mirjalili · 2022

Chapter 6

Model Evaluation and Hyperparameter Tuning

Chapter 10

Clustering Analysis

References and Suggested Reading

Reading

*Machine Learning with PyTorch and
Scikit-Learn*

Raschka, Liu and Mirjalili · 2022

Chapter 3: A Tour of Machine Learning
Classifiers Using Scikit-Learn

Decision Trees, Random Forests and kNN