UNIVERSITY OF CAMBRIDGE

Machine Learning for Chemists

– Decision Trees & Random Forests –

---

26 February 2024

Outline

UNIVERSITY OF CAMBRIDGE

Introduction to Decision Trees

Regression with Decision Trees

The “big idea” behind Random Forests

Examples of RFs used in the wild

Live Demo!

A Familiar Example: Decision Trees

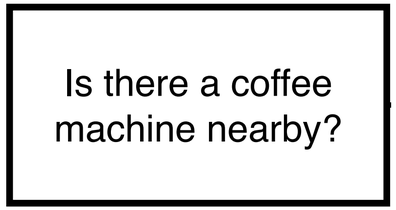
UNIVERSITY OF CAMBRIDGE

Should I get some coffee flowchart?

A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

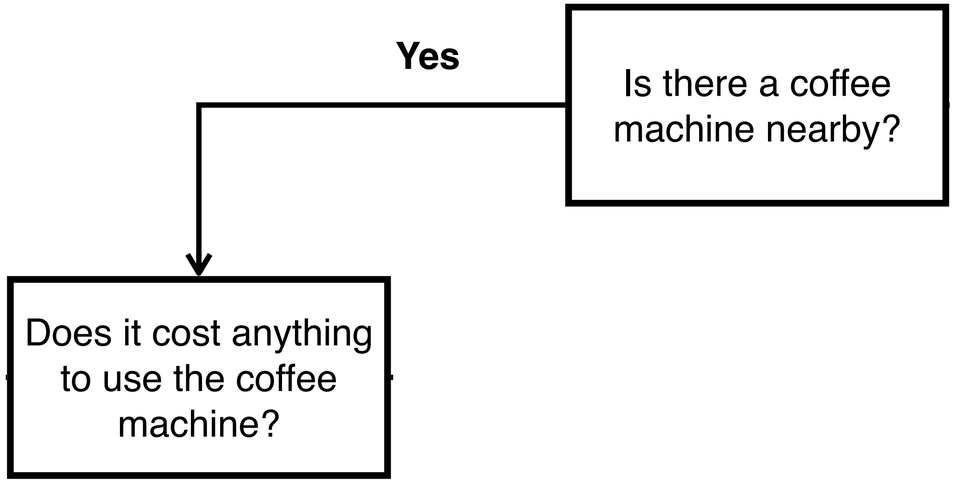
Should I get some coffee flowchart?



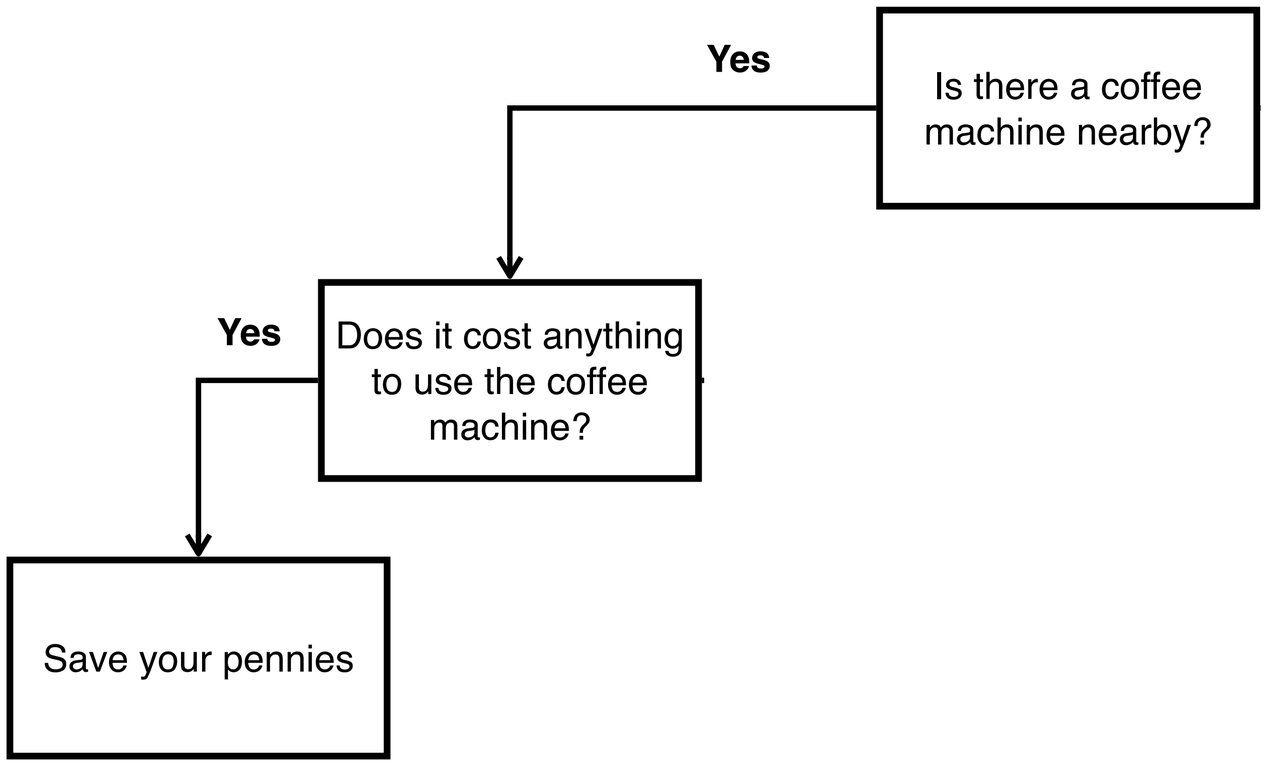
A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

Should I get some coffee flowchart?



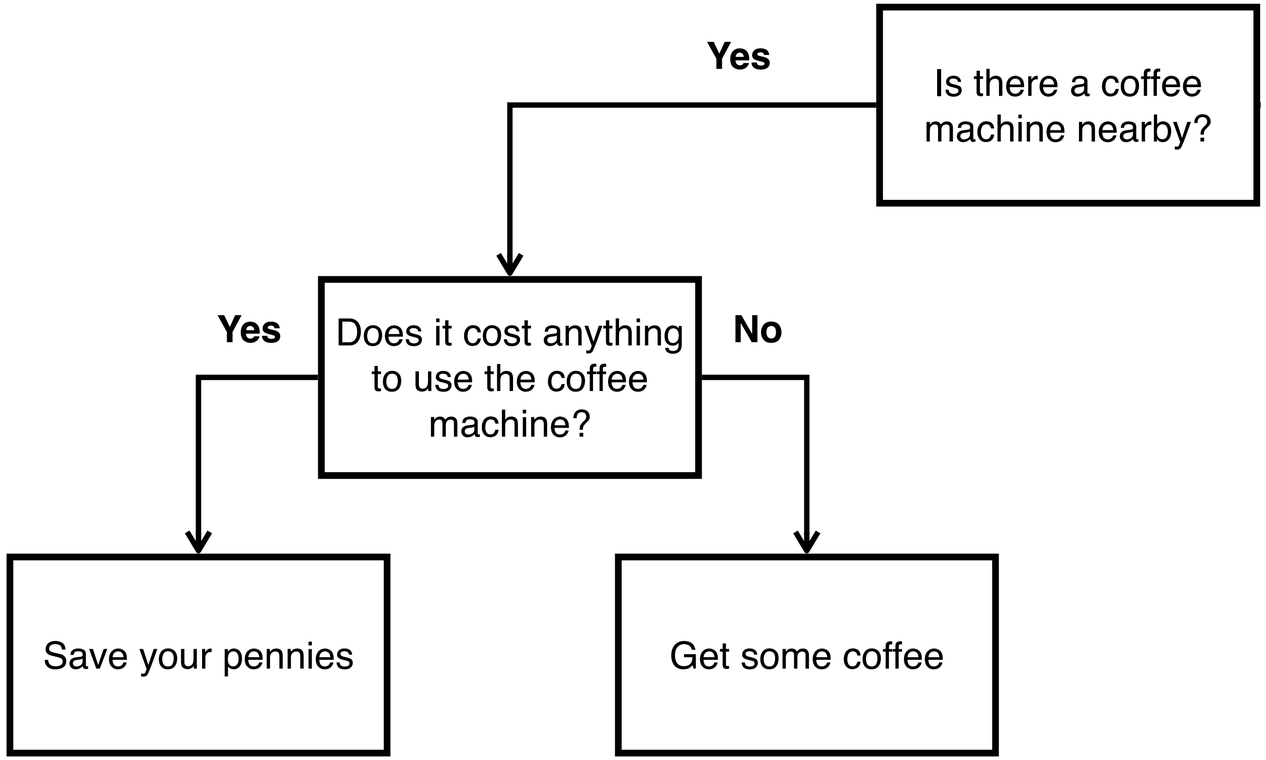
A Familiar Example: Decision Trees



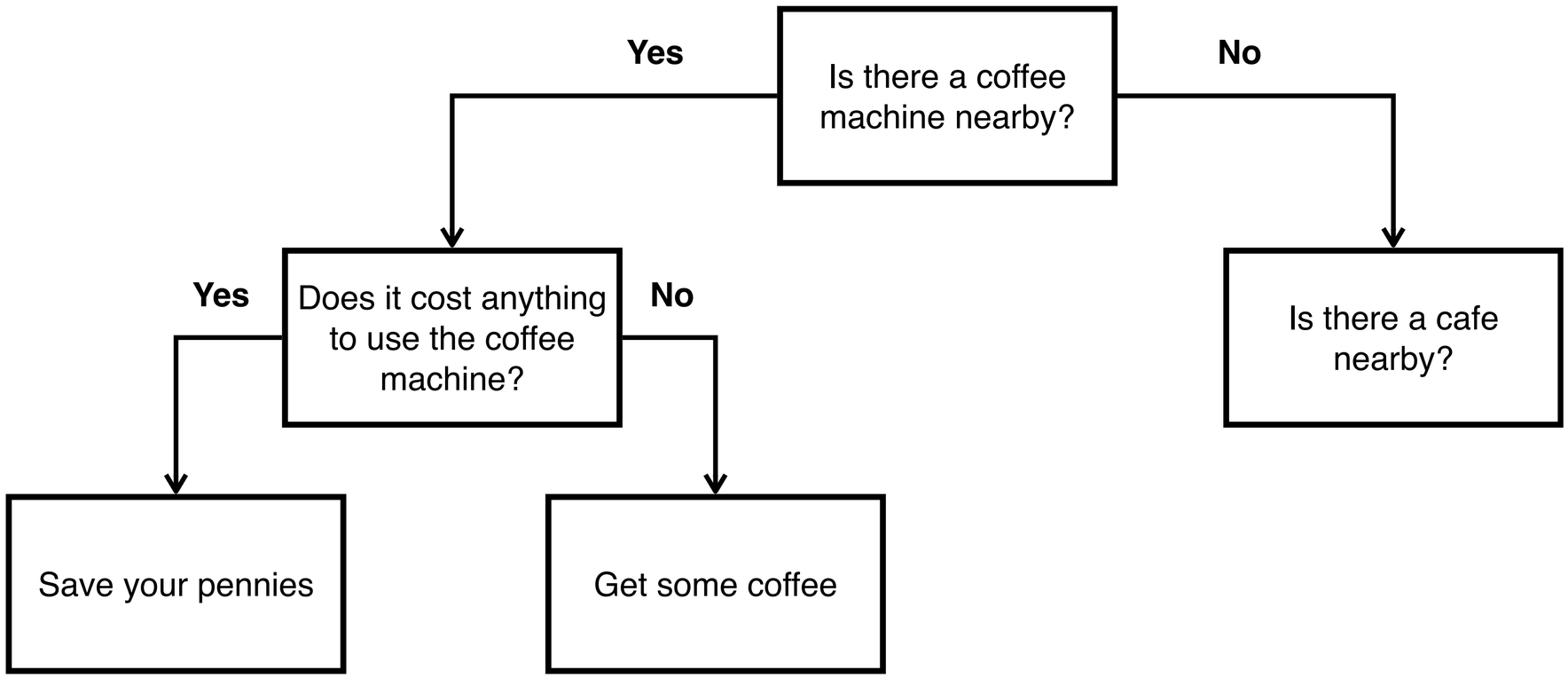
A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

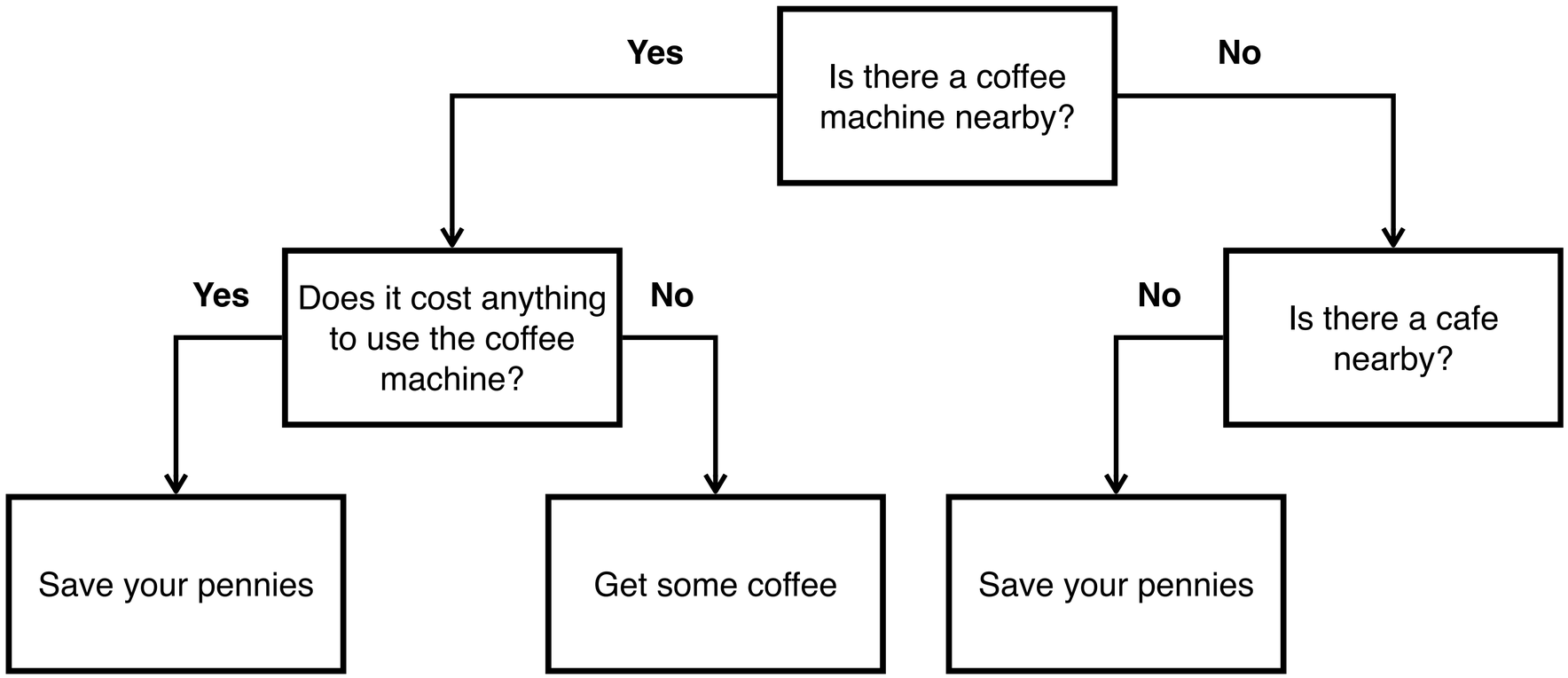
Should I get some coffee flowchart?



A Familiar Example: Decision Trees



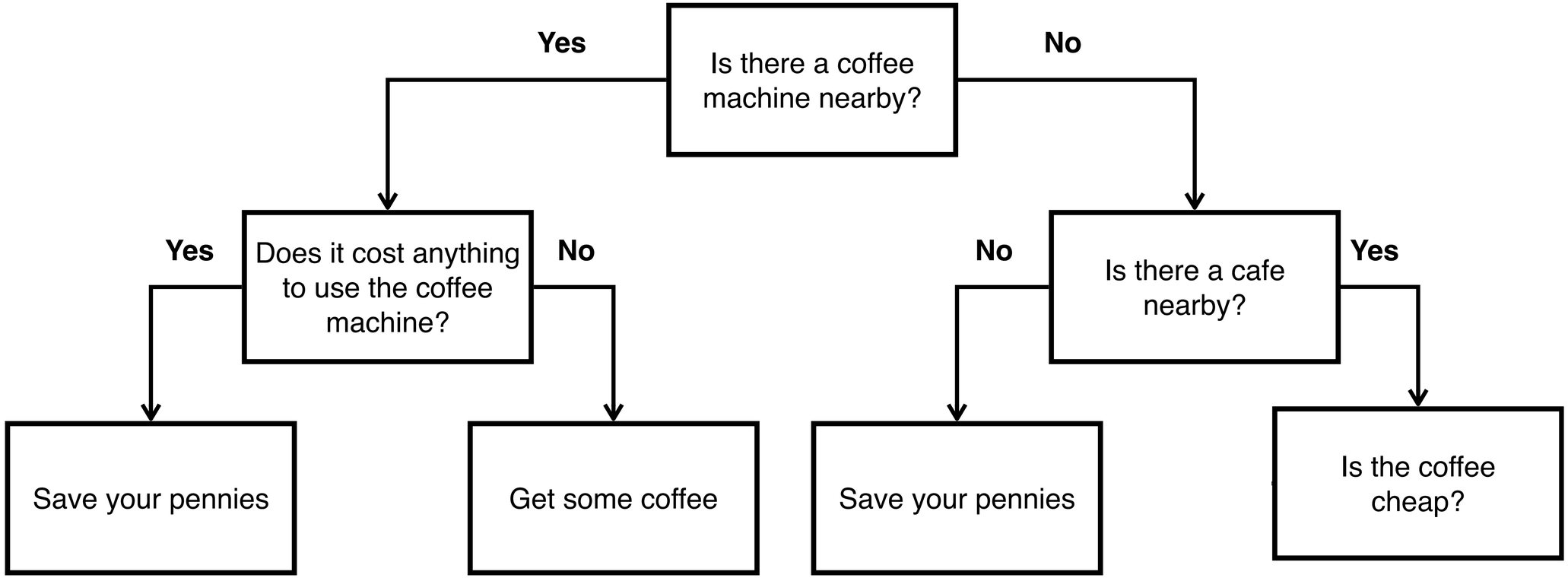
A Familiar Example: Decision Trees



A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

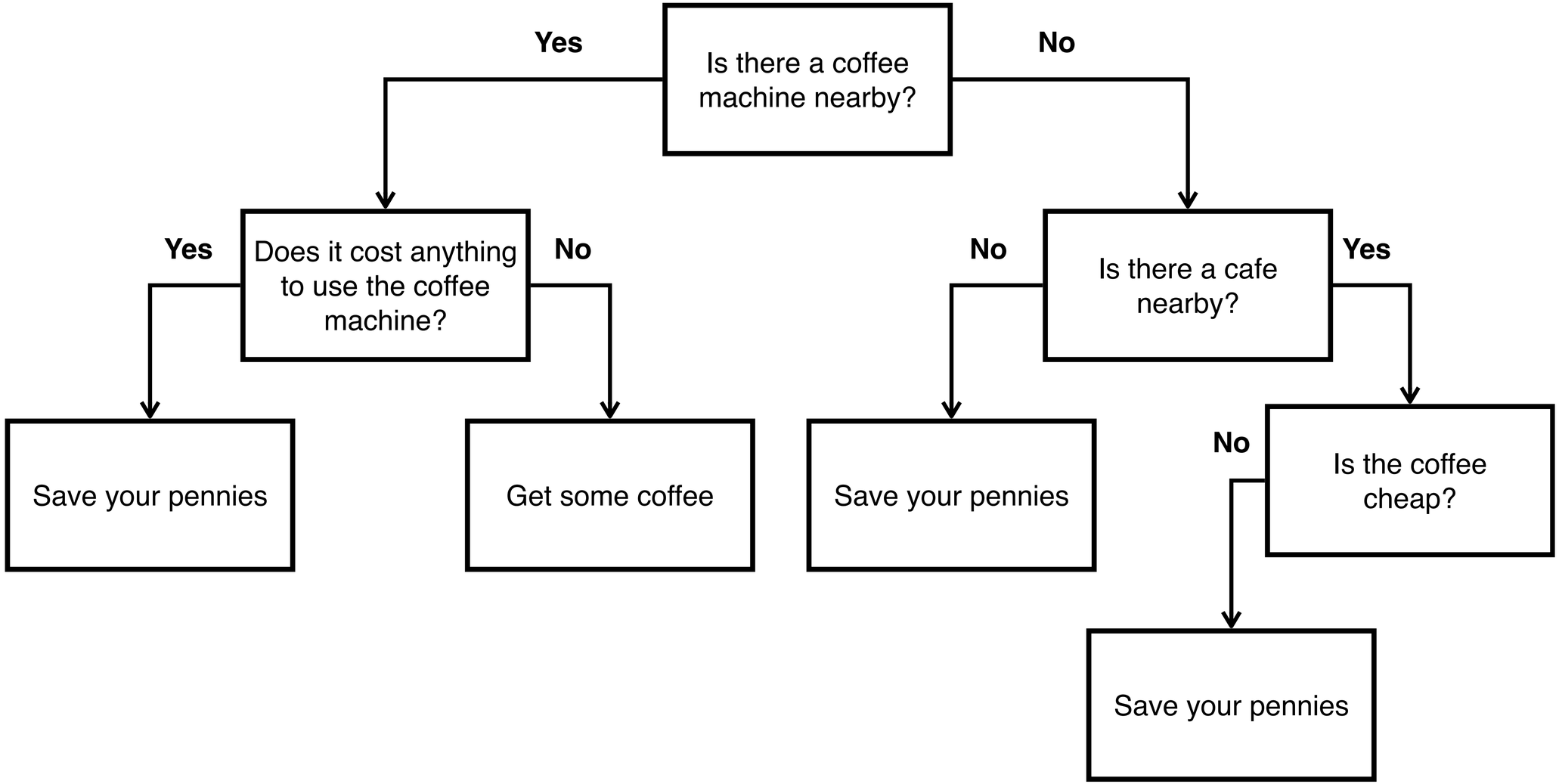
Should I get some coffee flowchart?



A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

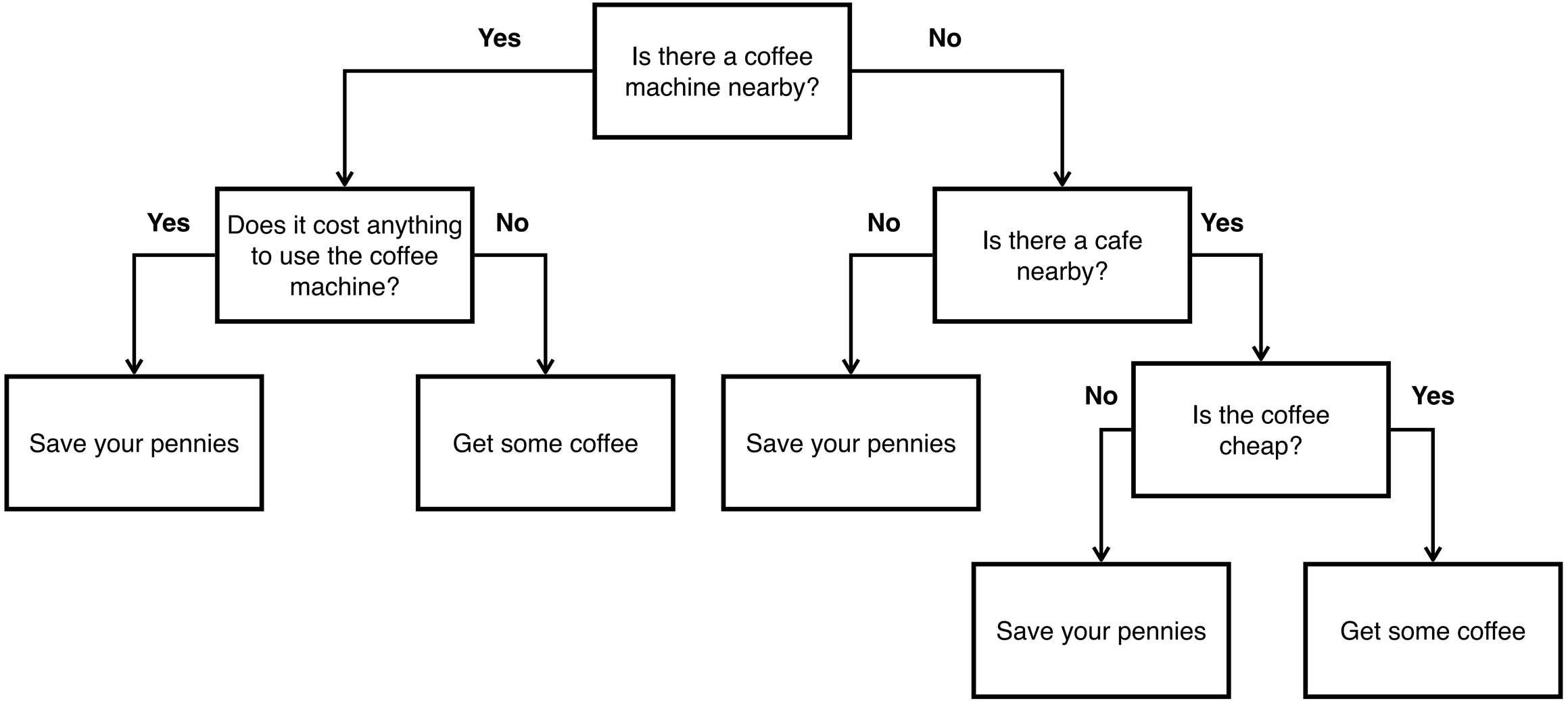
Should I get some coffee flowchart?



A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

Should I get some coffee flowchart?



A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

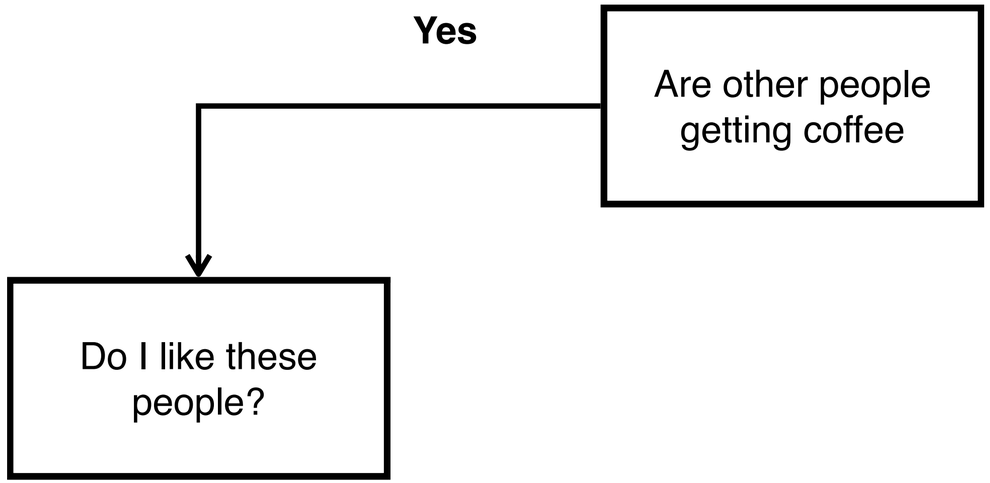
Should I get some coffee flowchart?

Are other people getting coffee

A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

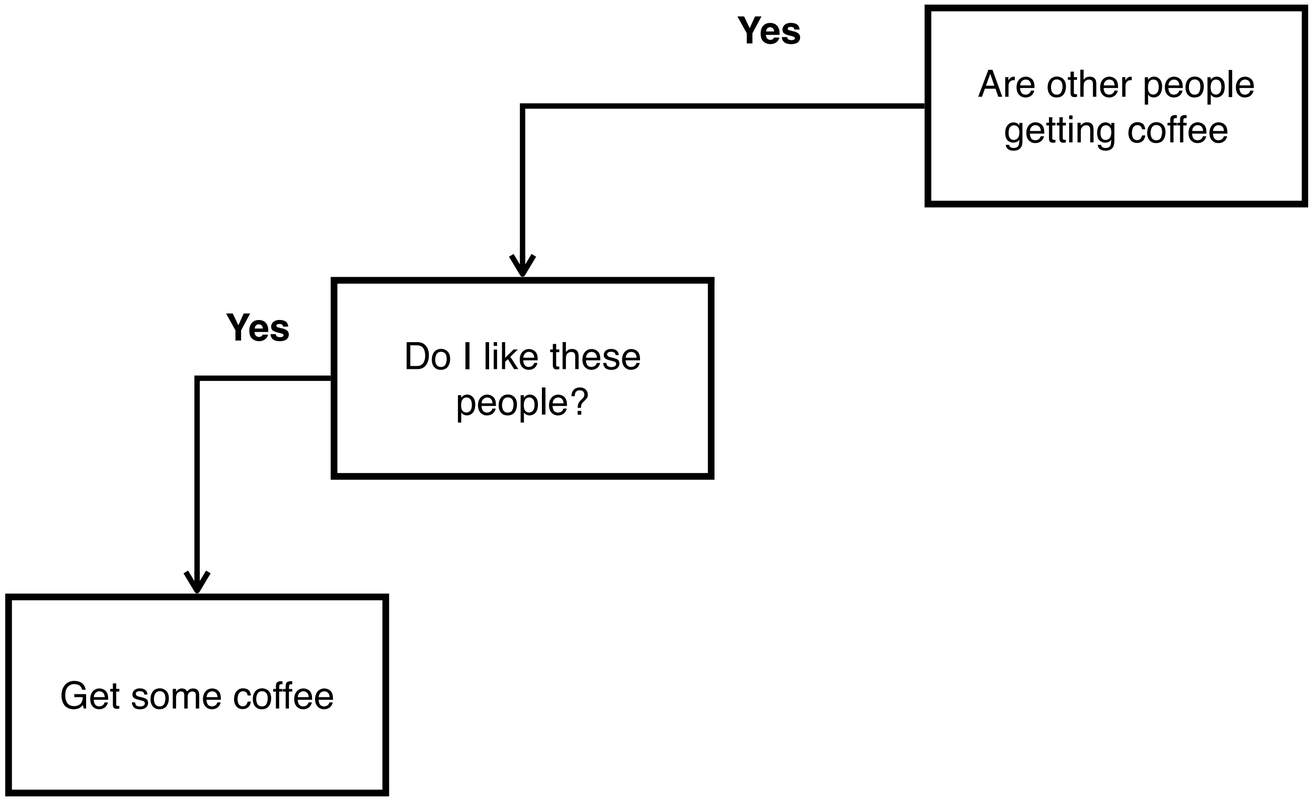
Should I get some coffee flowchart?



A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

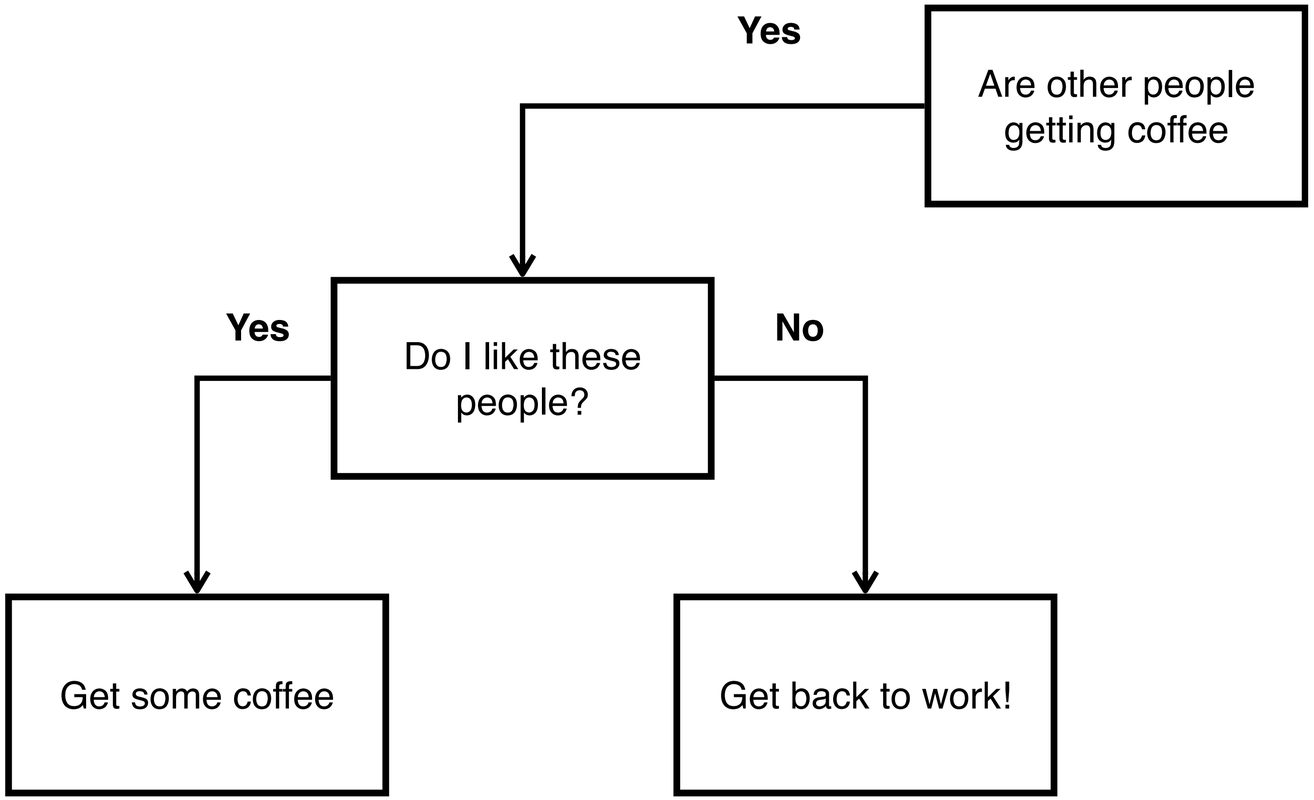
Should I get some coffee flowchart?



A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

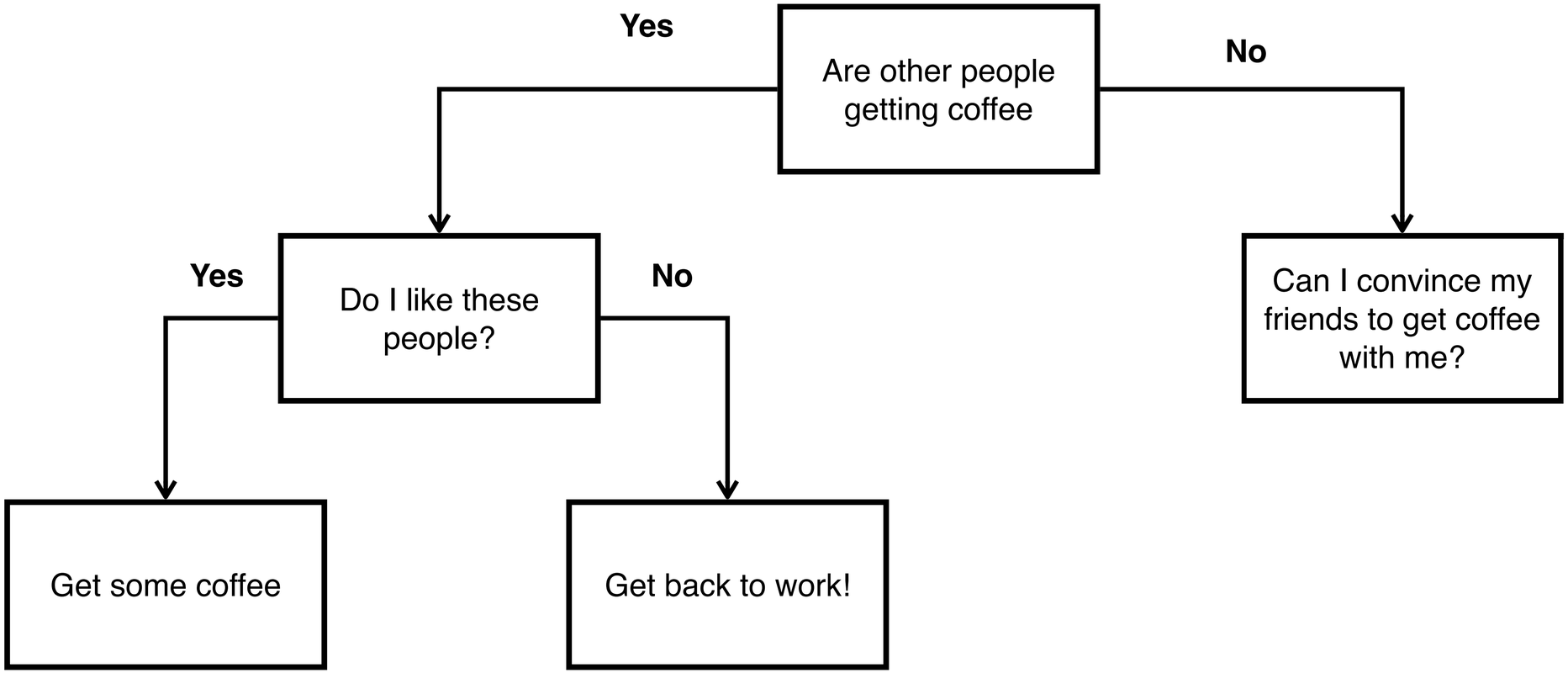
Should I get some coffee flowchart?



A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

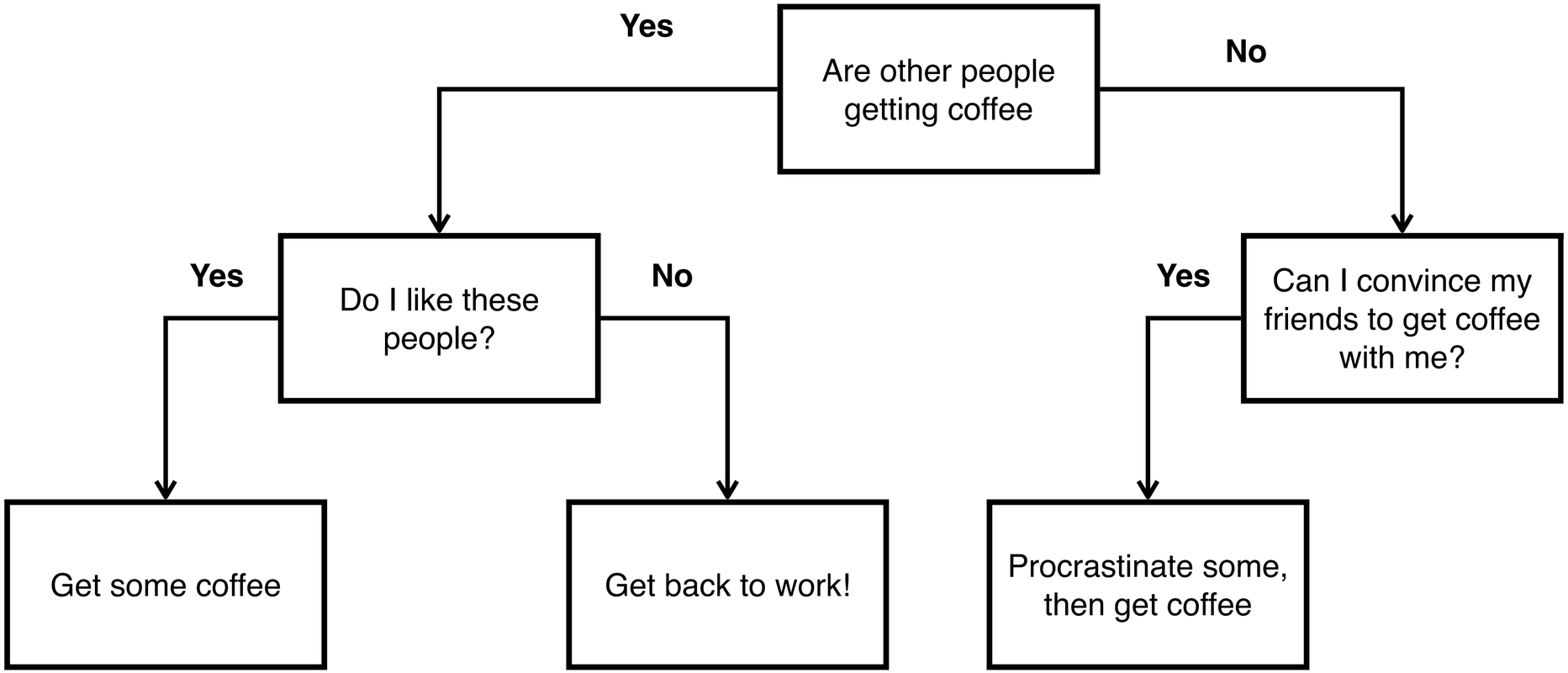
Should I get some coffee flowchart?



A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

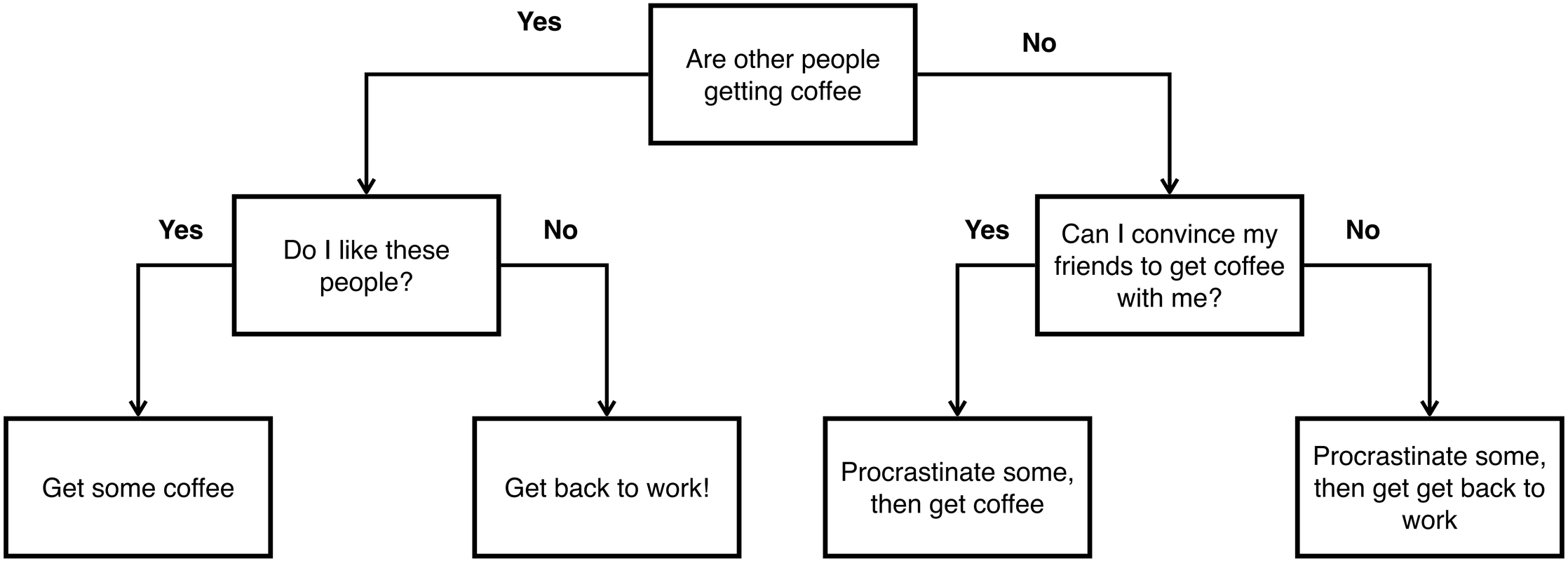
Should I get some coffee flowchart?



A Familiar Example: Decision Trees

UNIVERSITY OF CAMBRIDGE

Should I get some coffee flowchart?



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

Variables = Branching points

Prediction = End points

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No

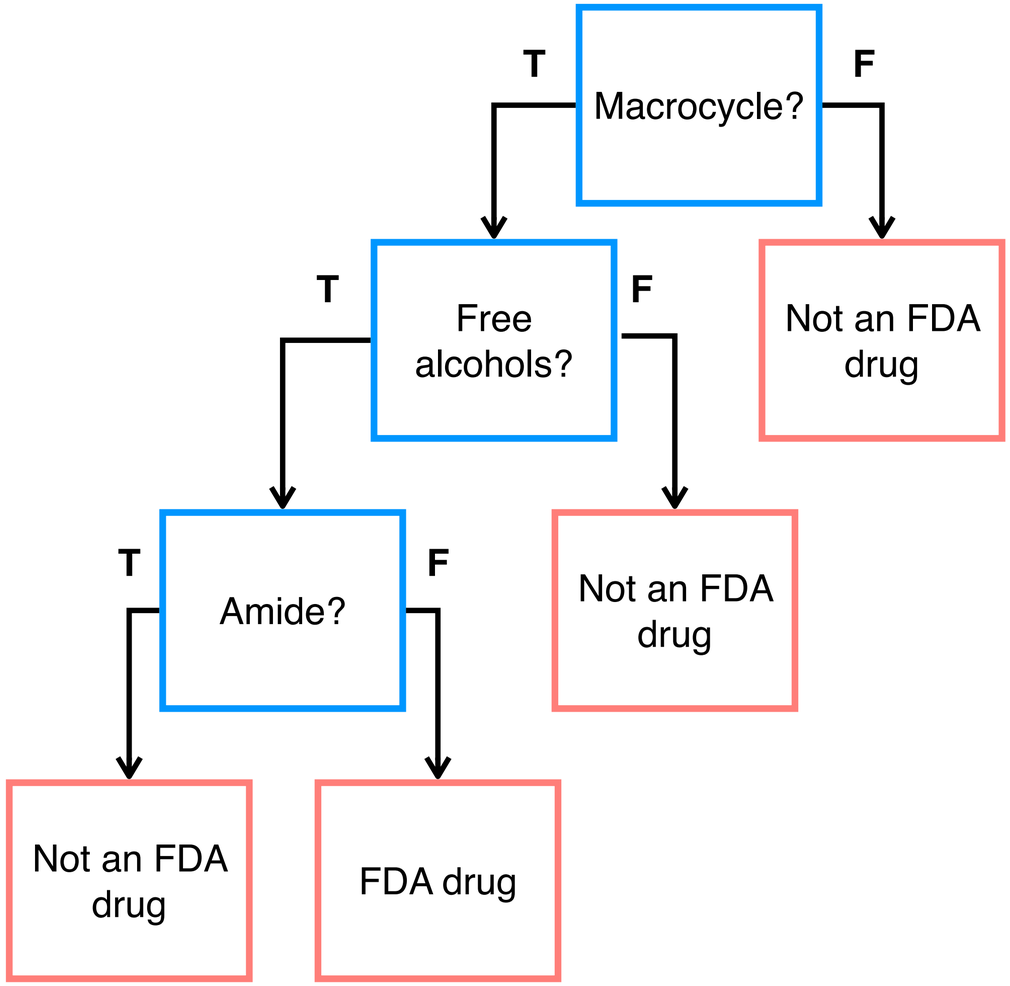
Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

Variables = Branching points

Prediction = End points

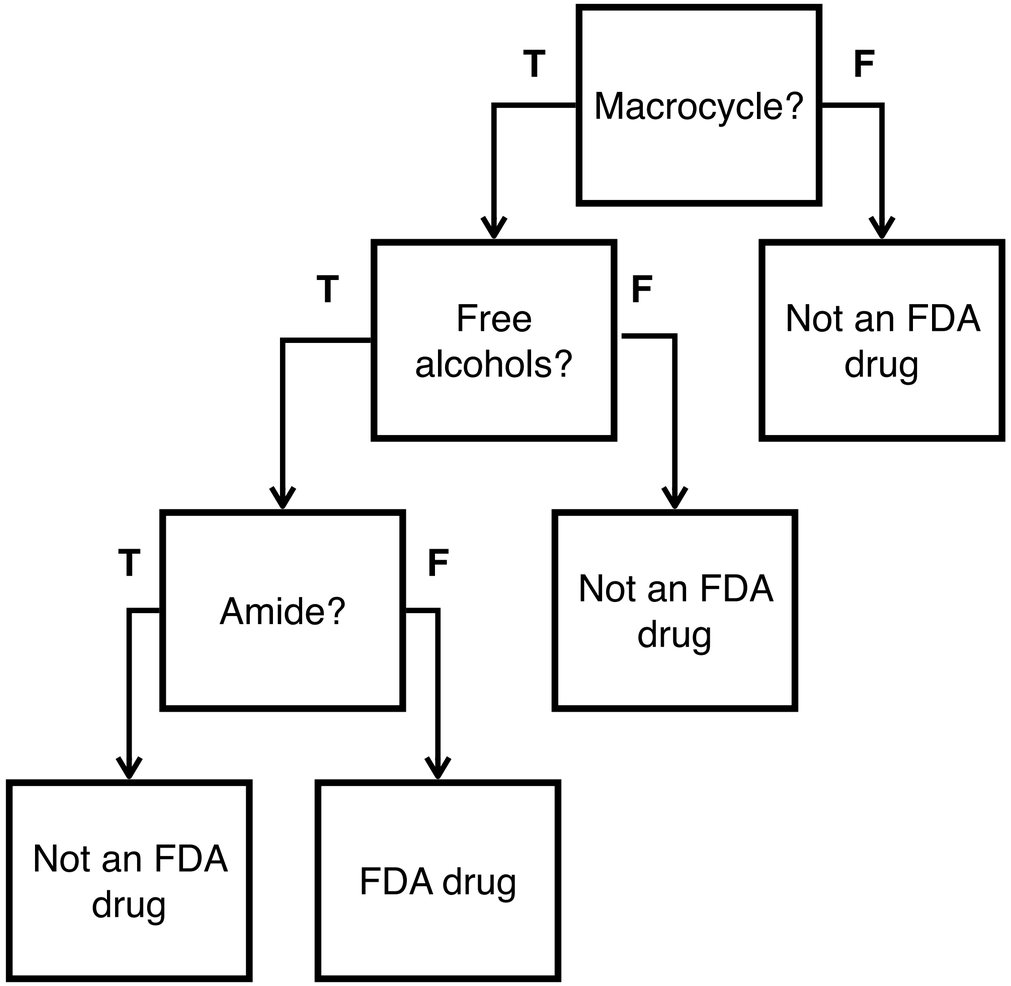
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

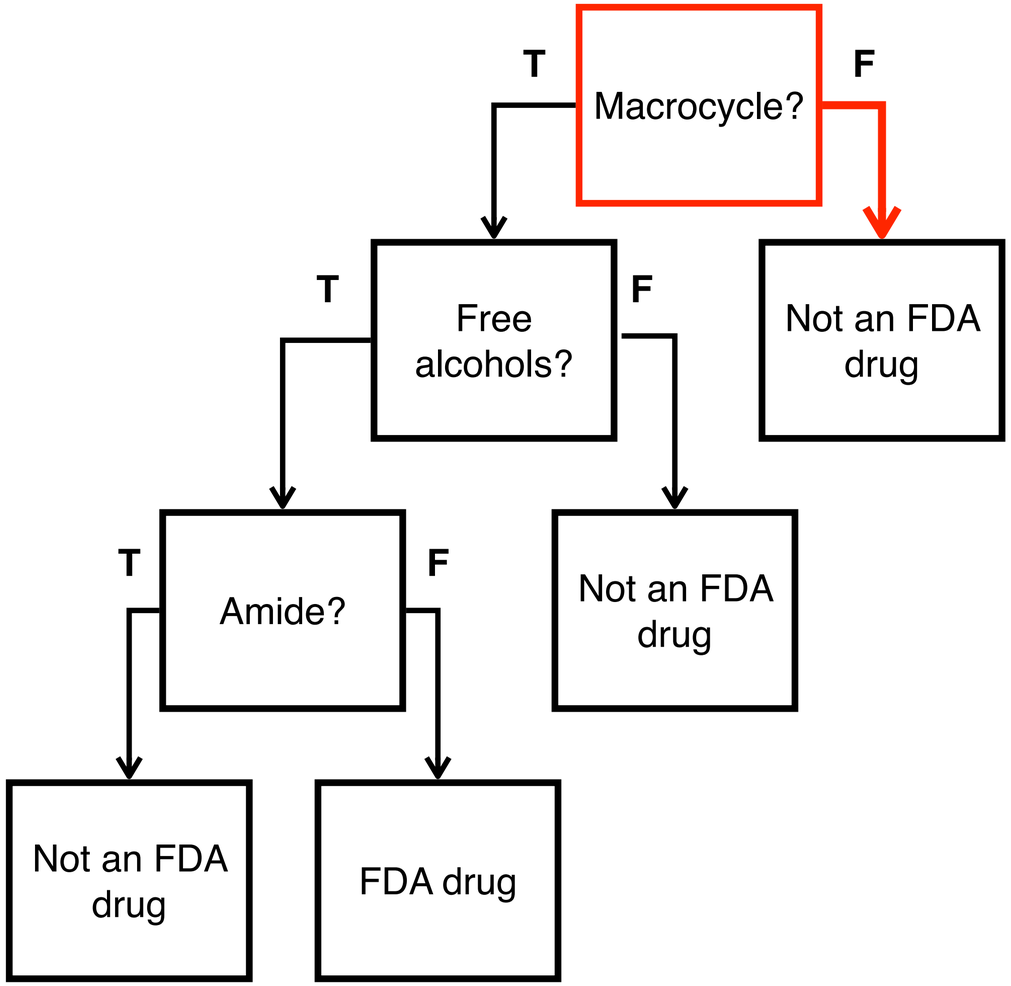
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

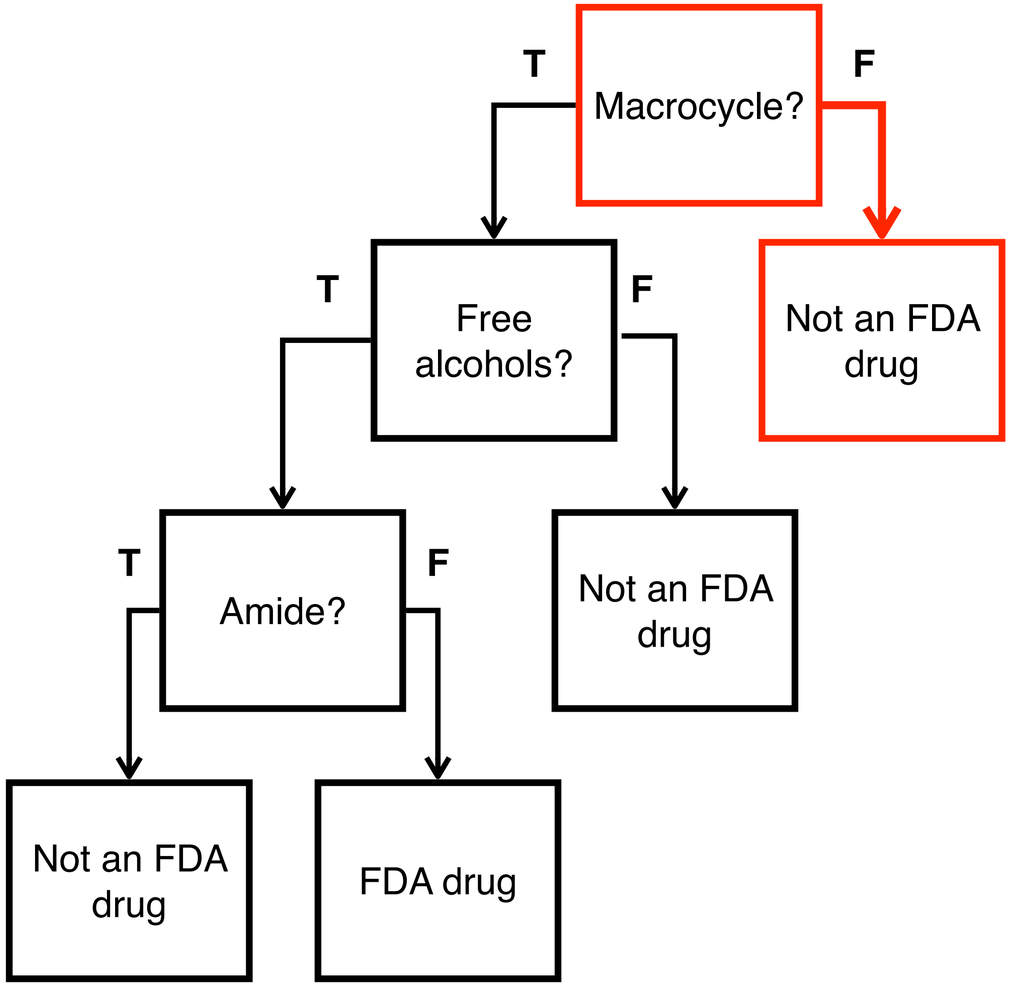
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

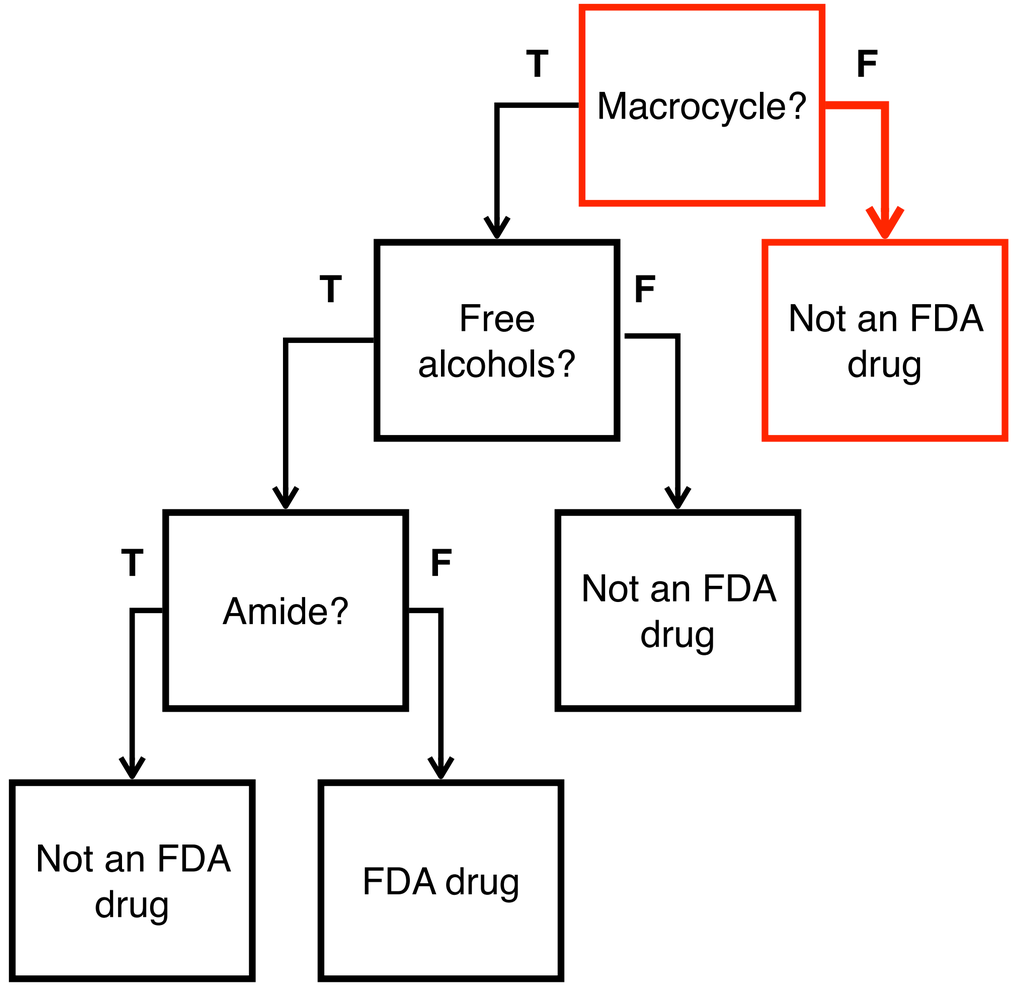
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

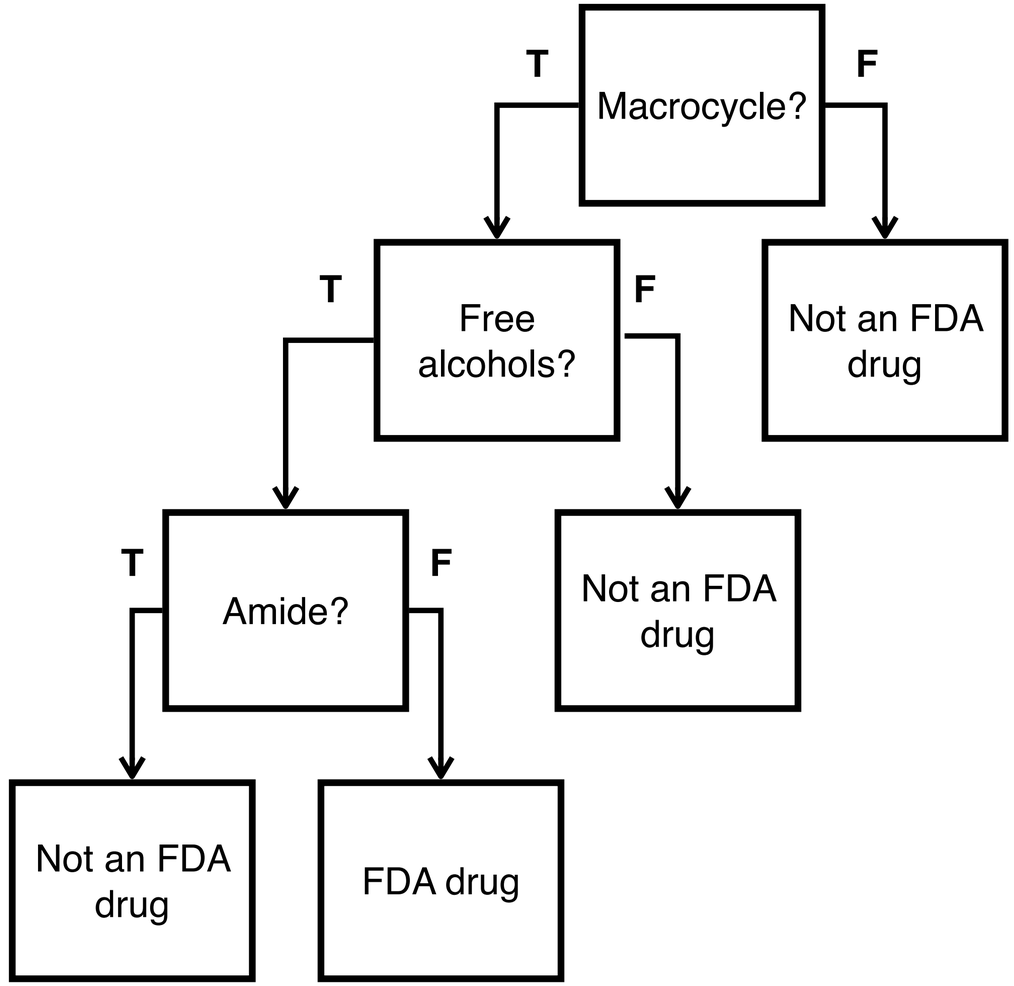
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

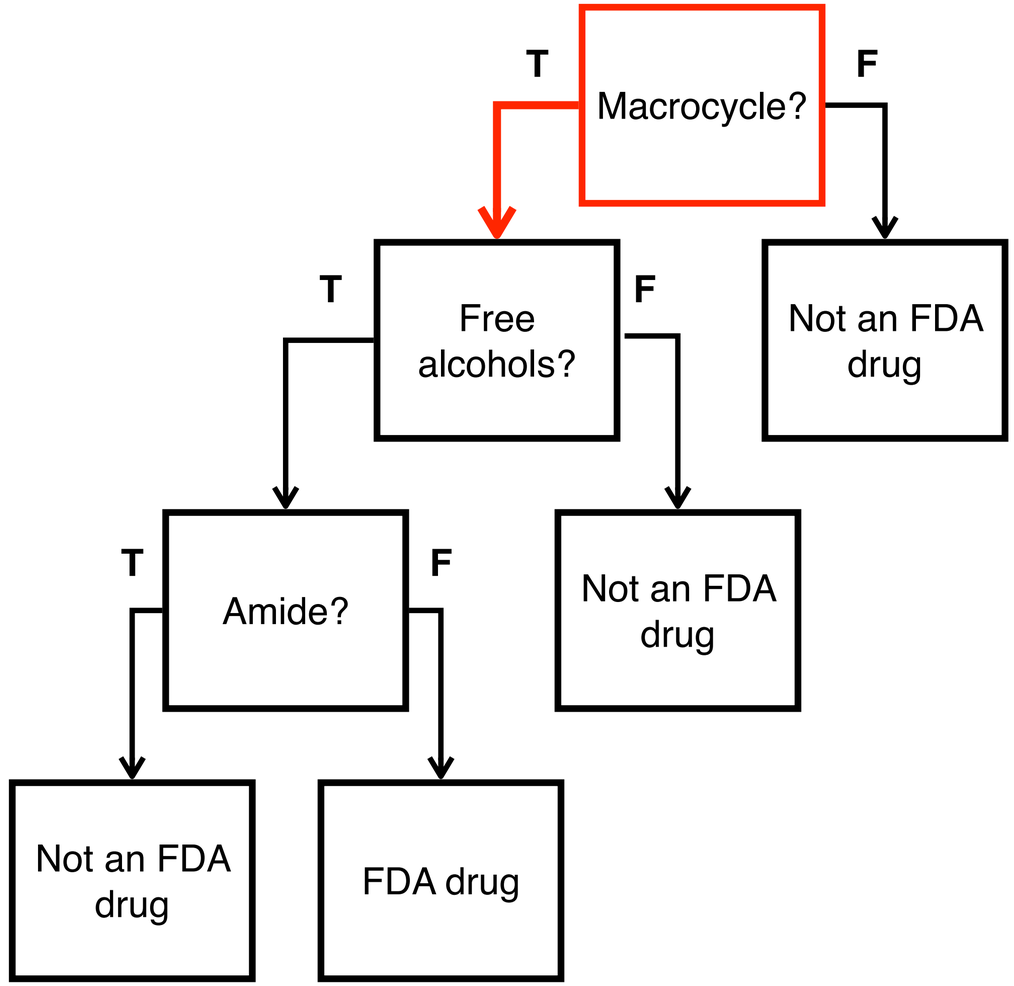
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

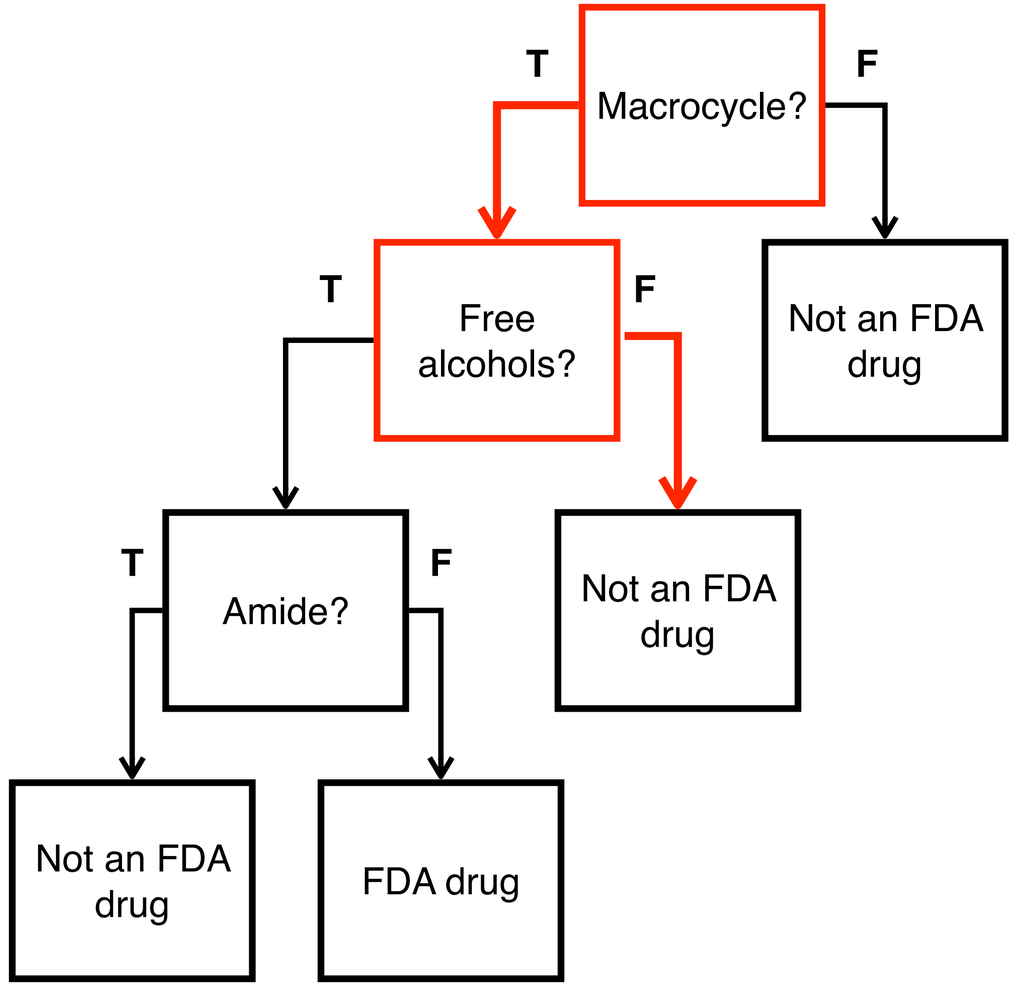
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

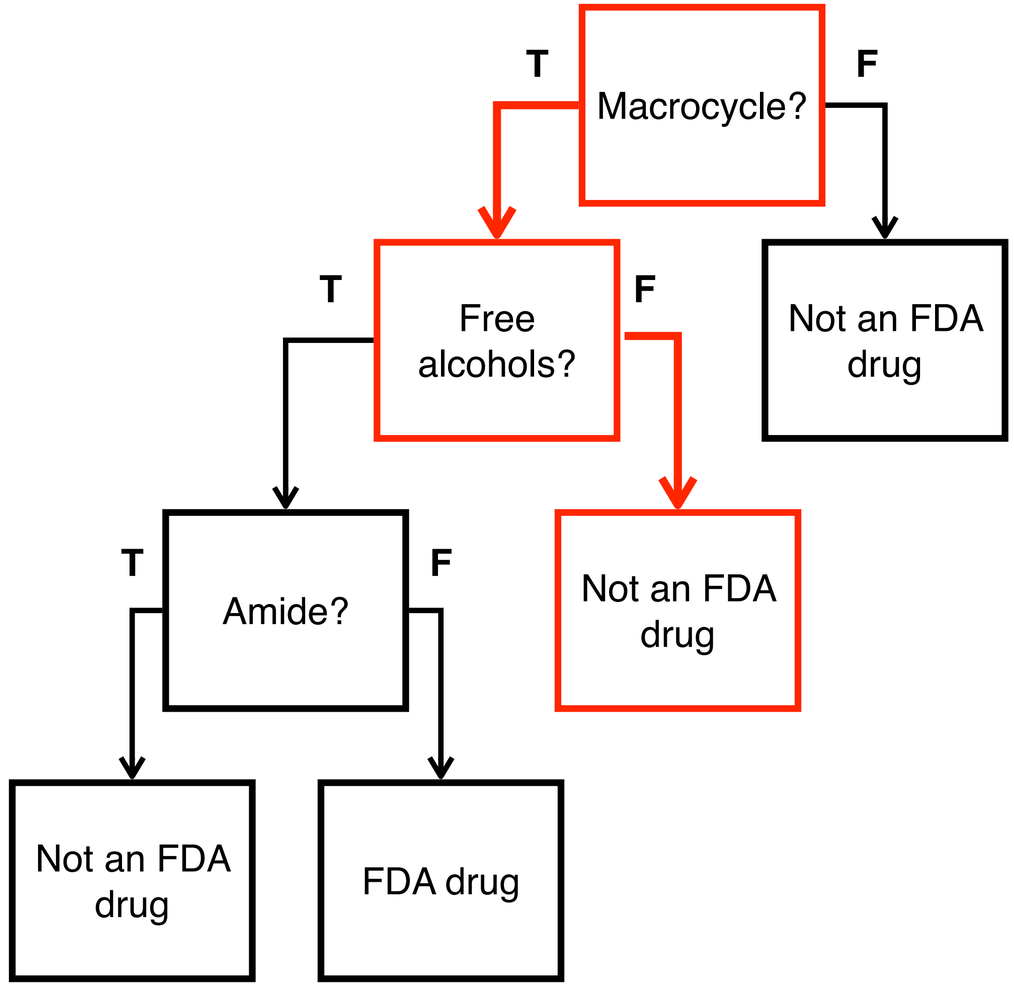
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

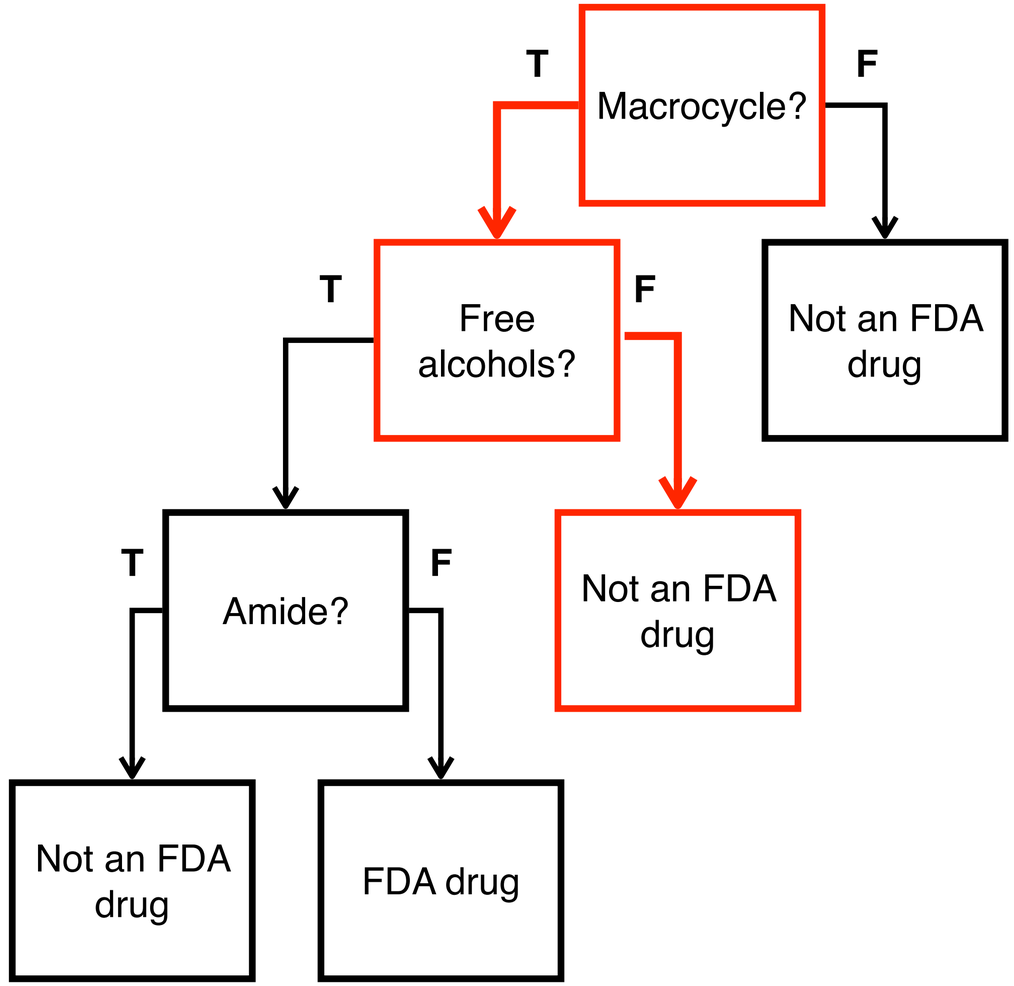
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

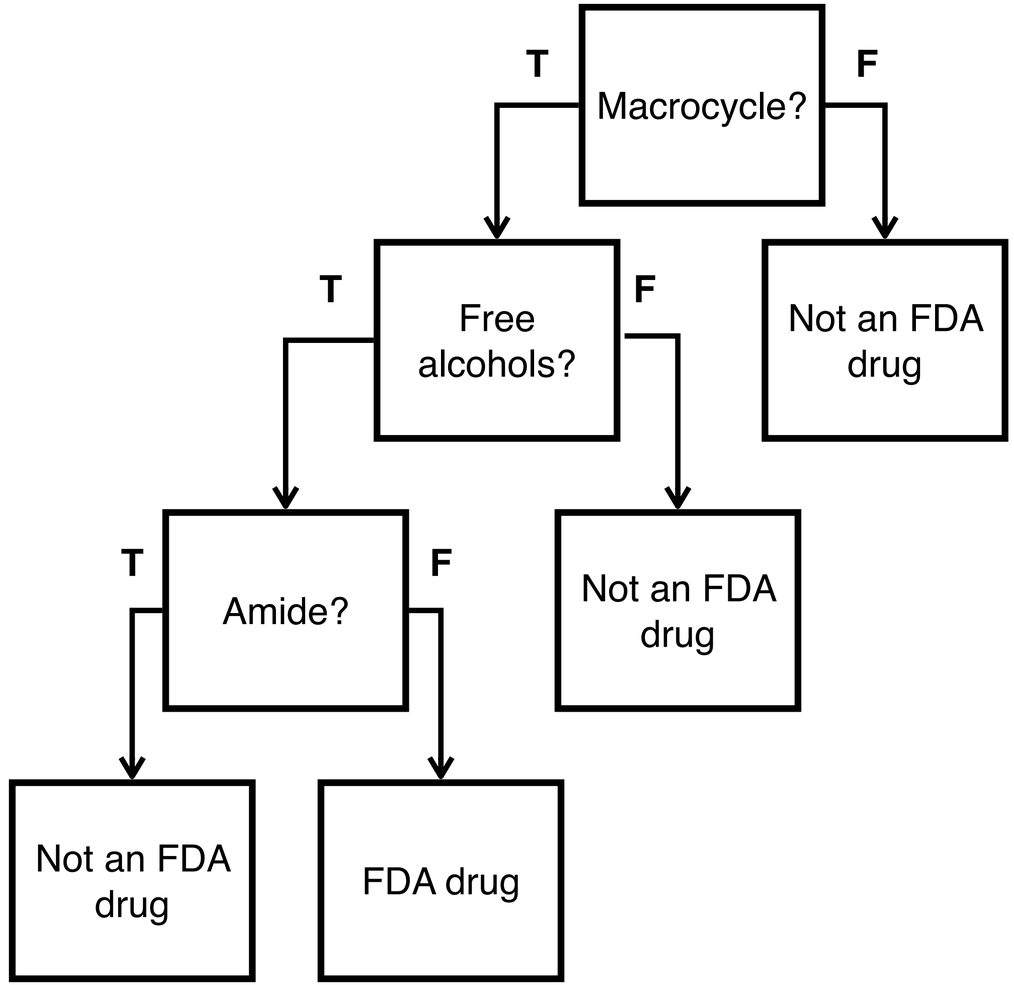
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

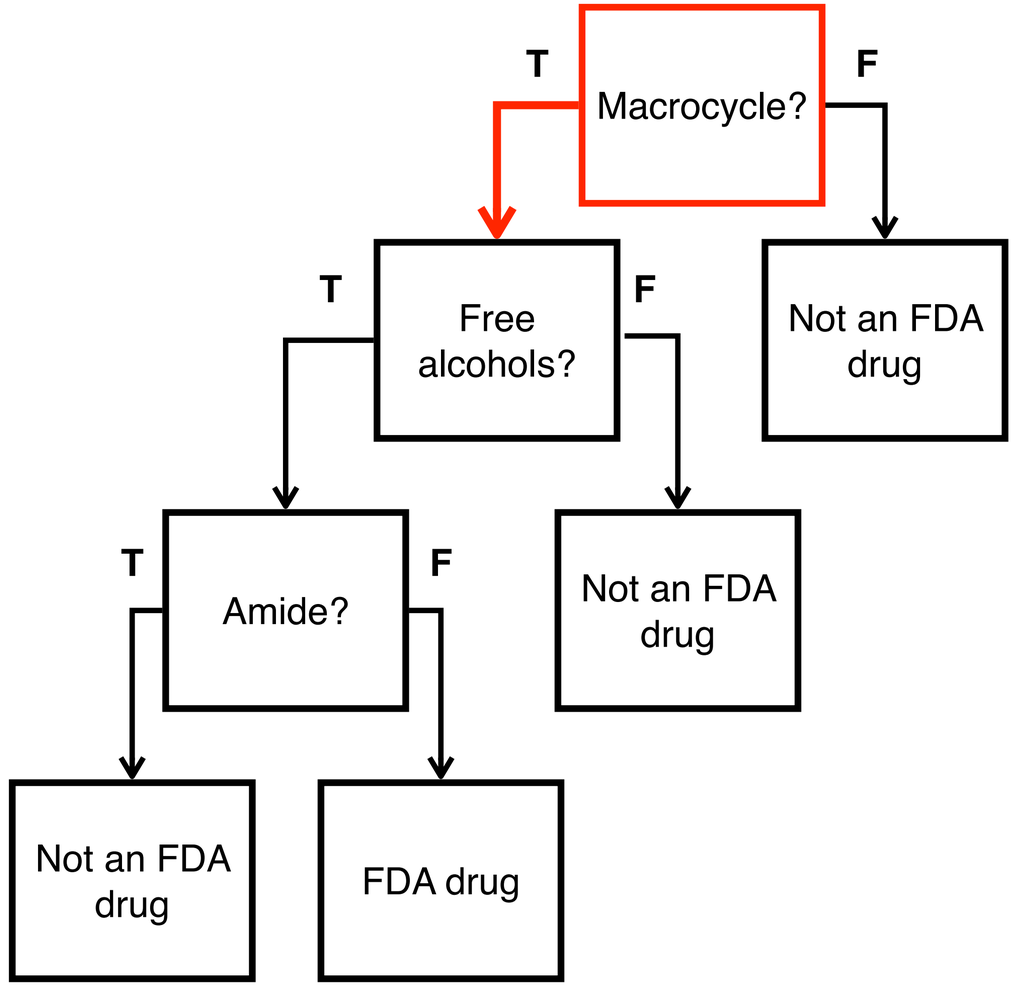
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

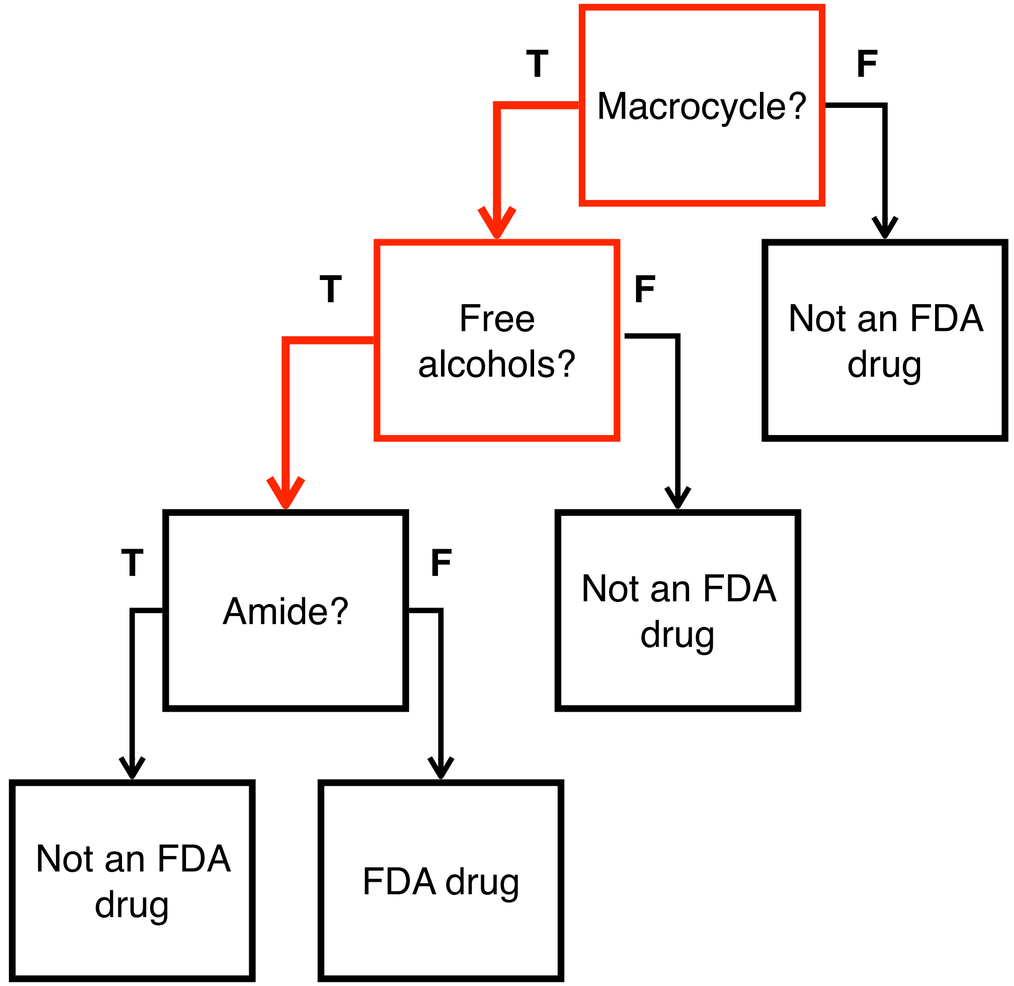
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

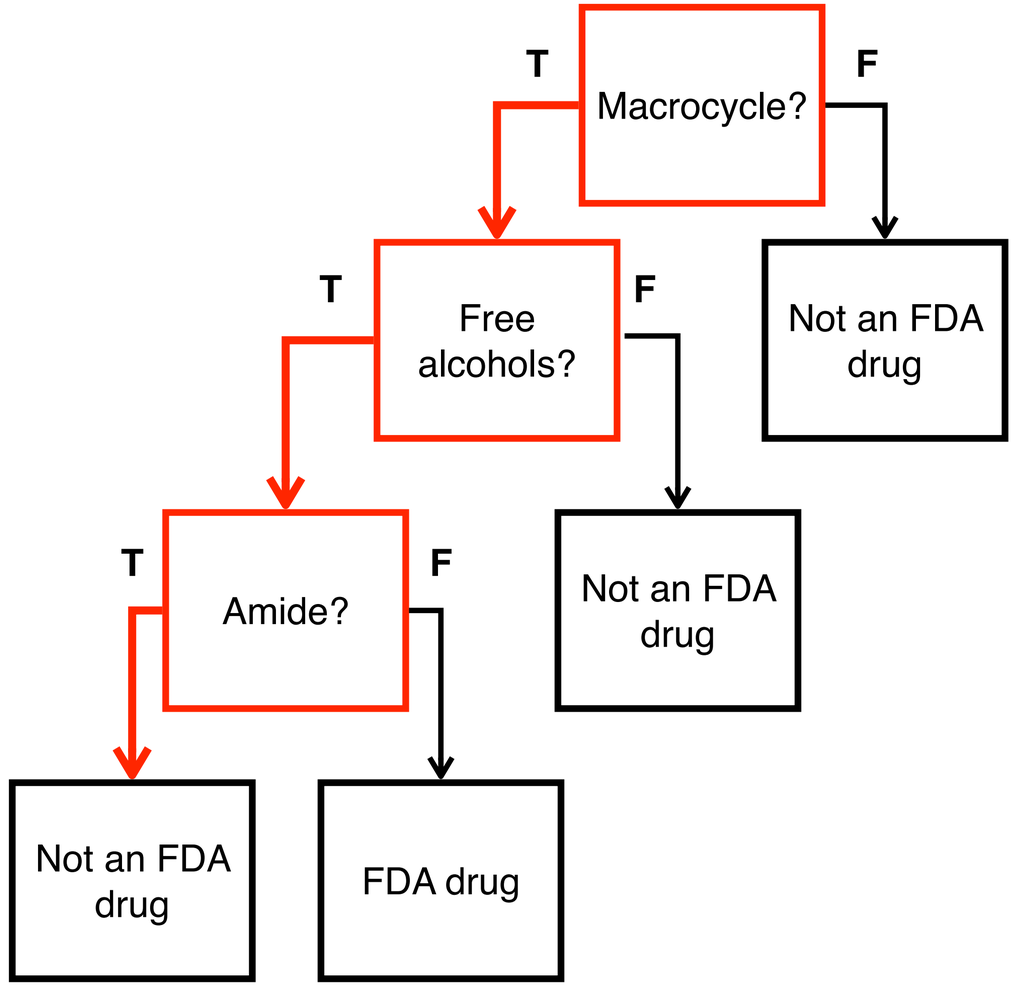
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

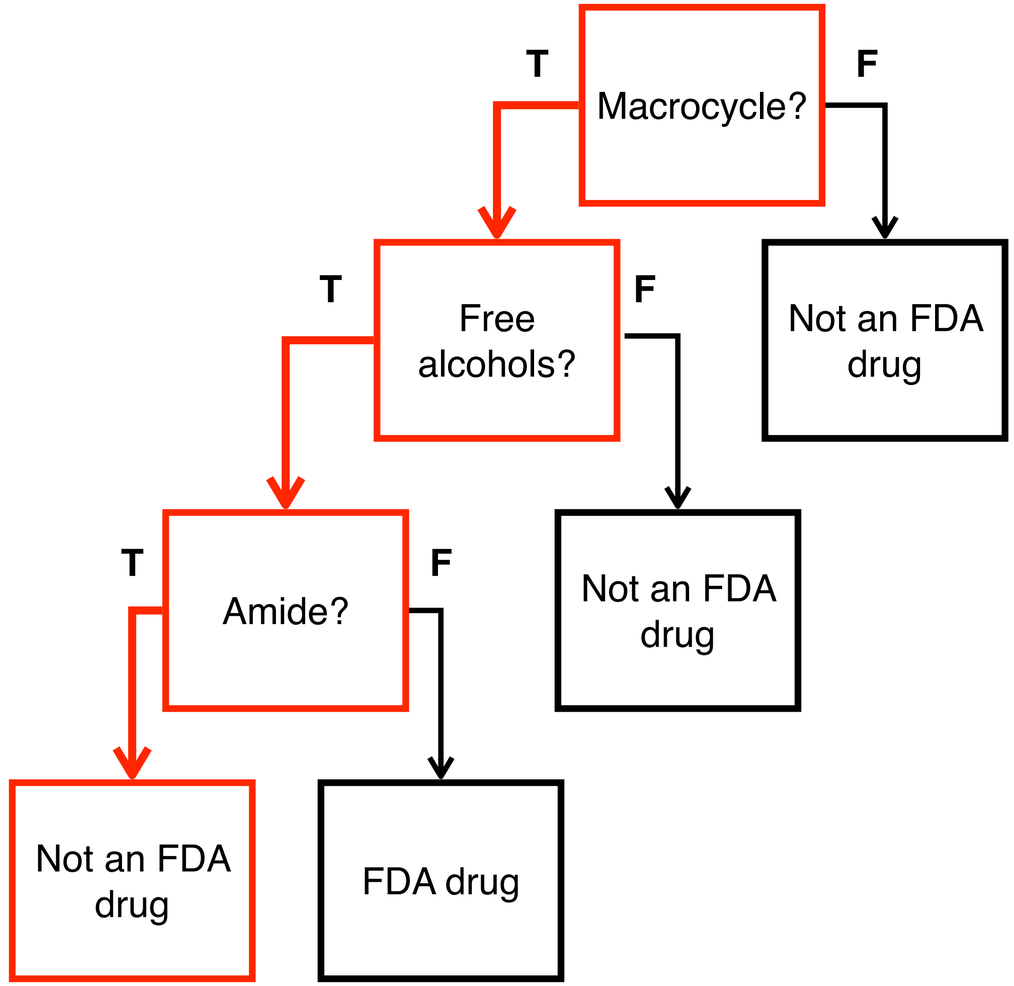
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

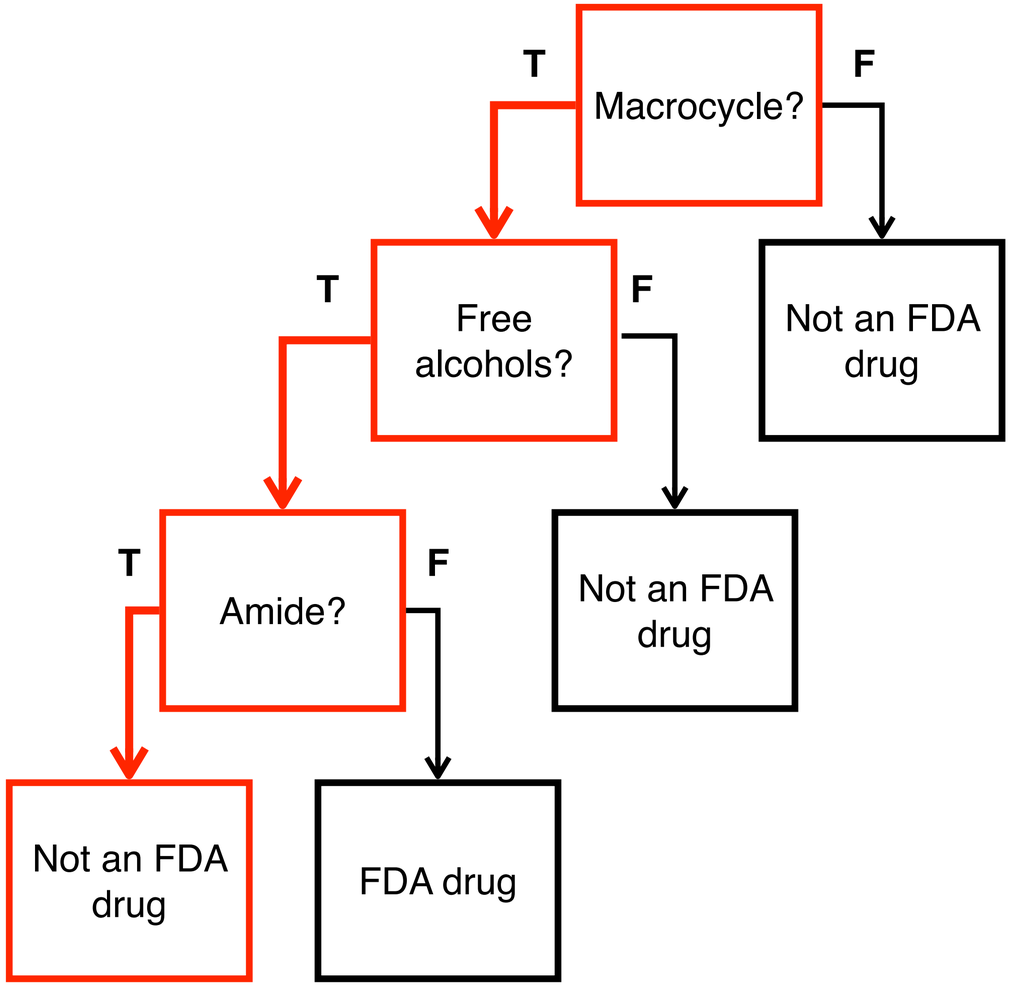
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

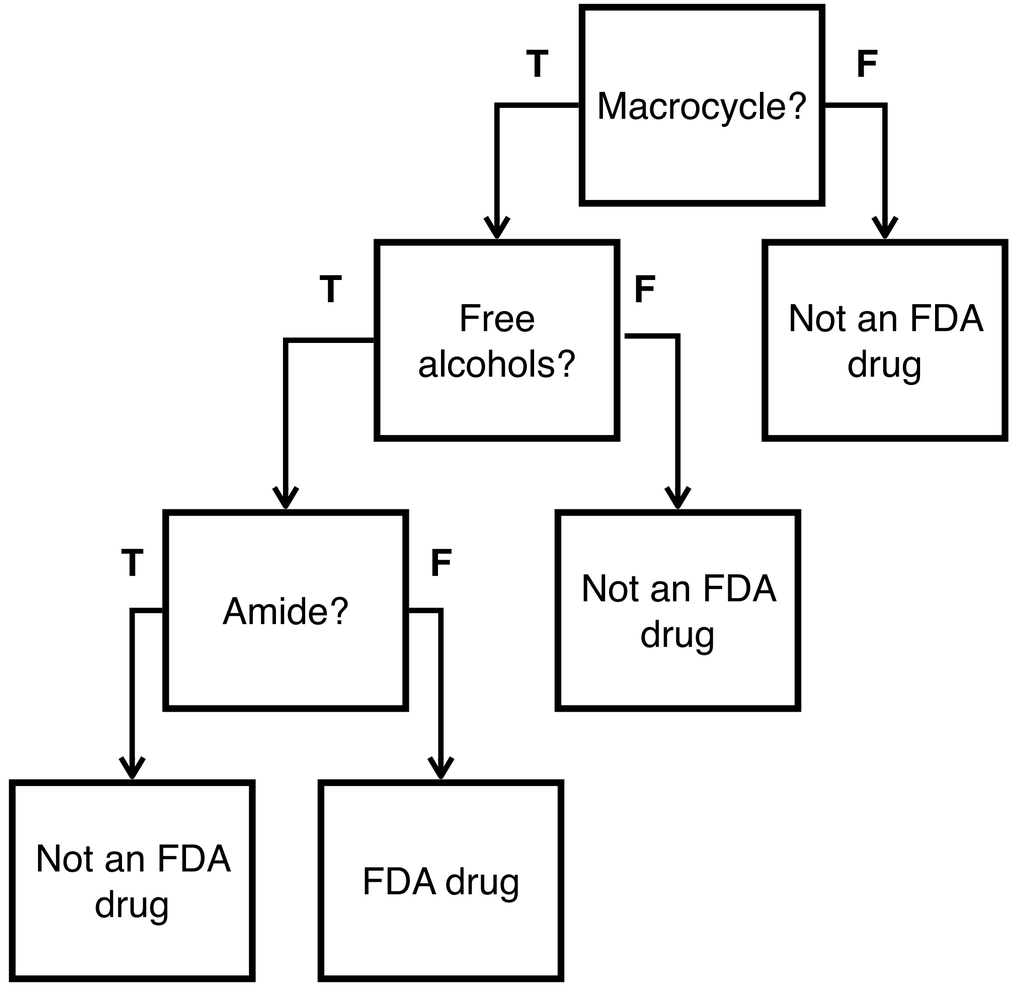
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

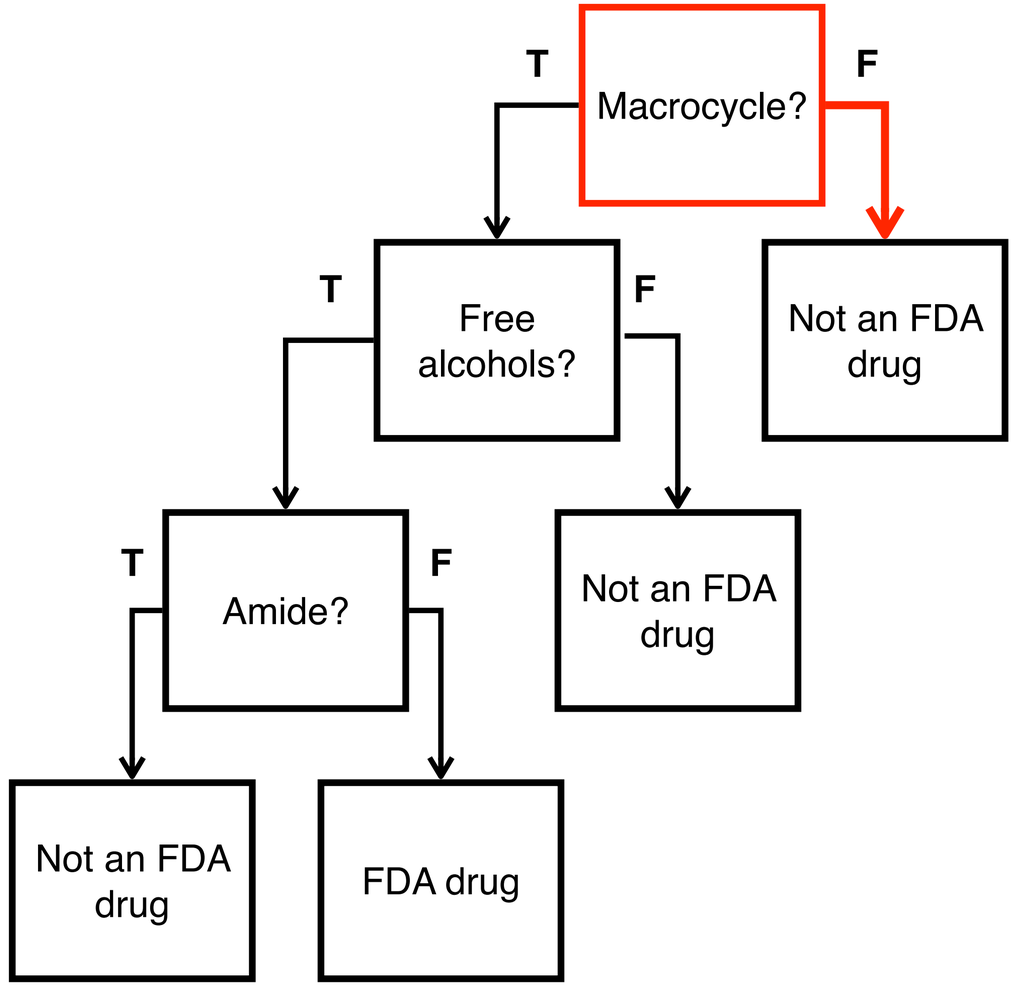
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

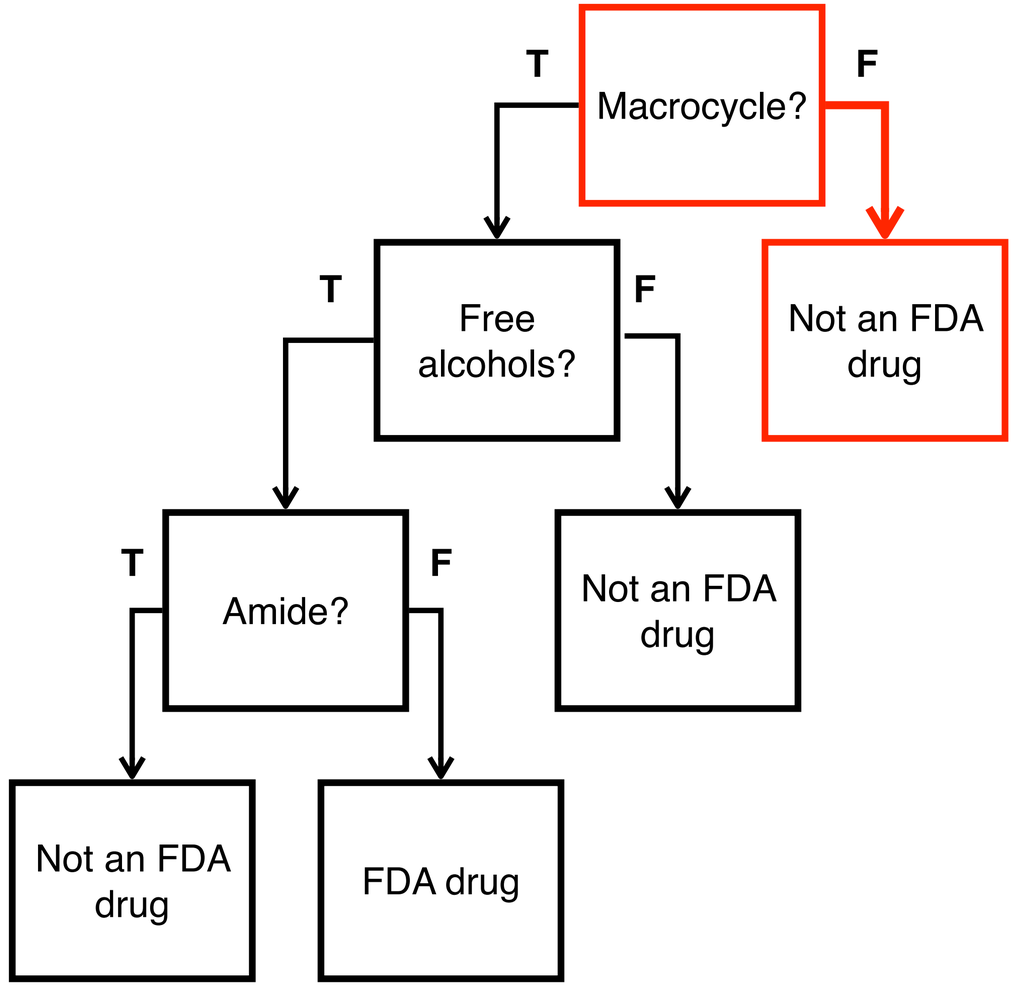
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

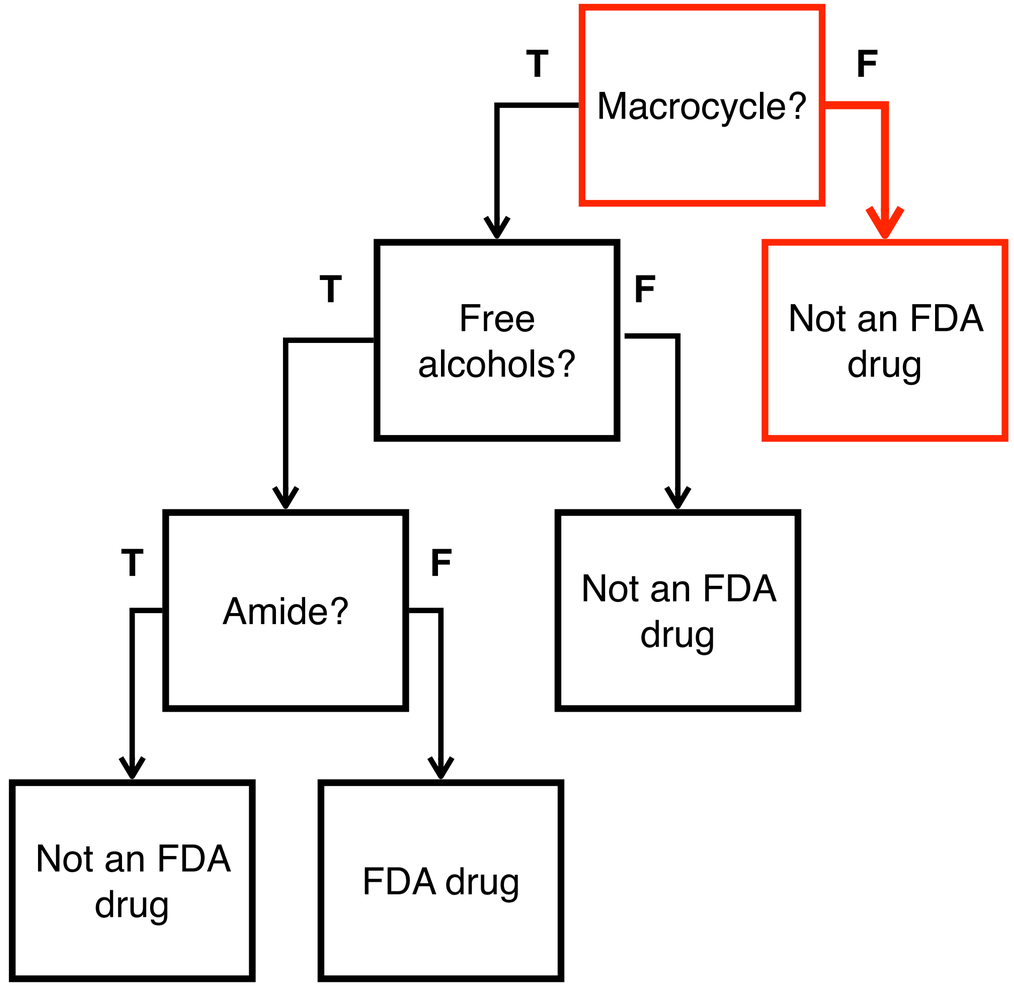
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

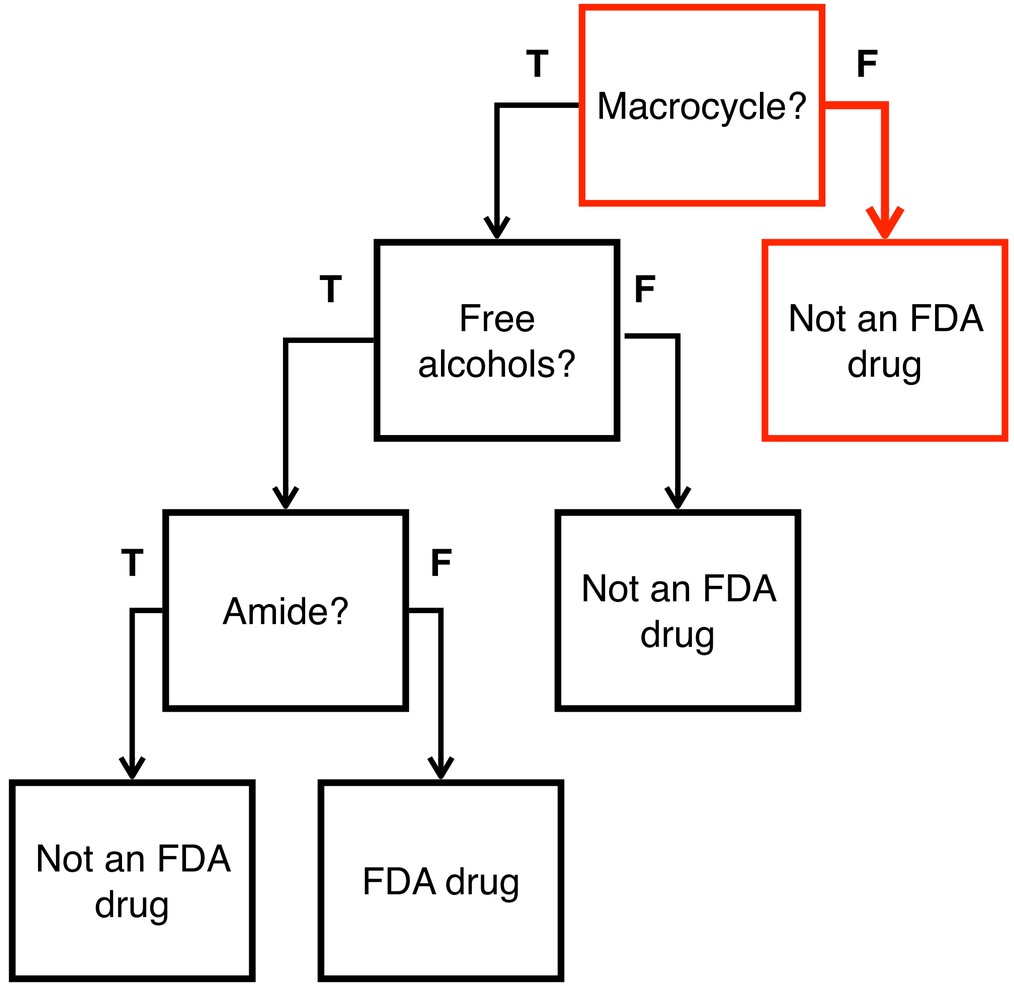
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

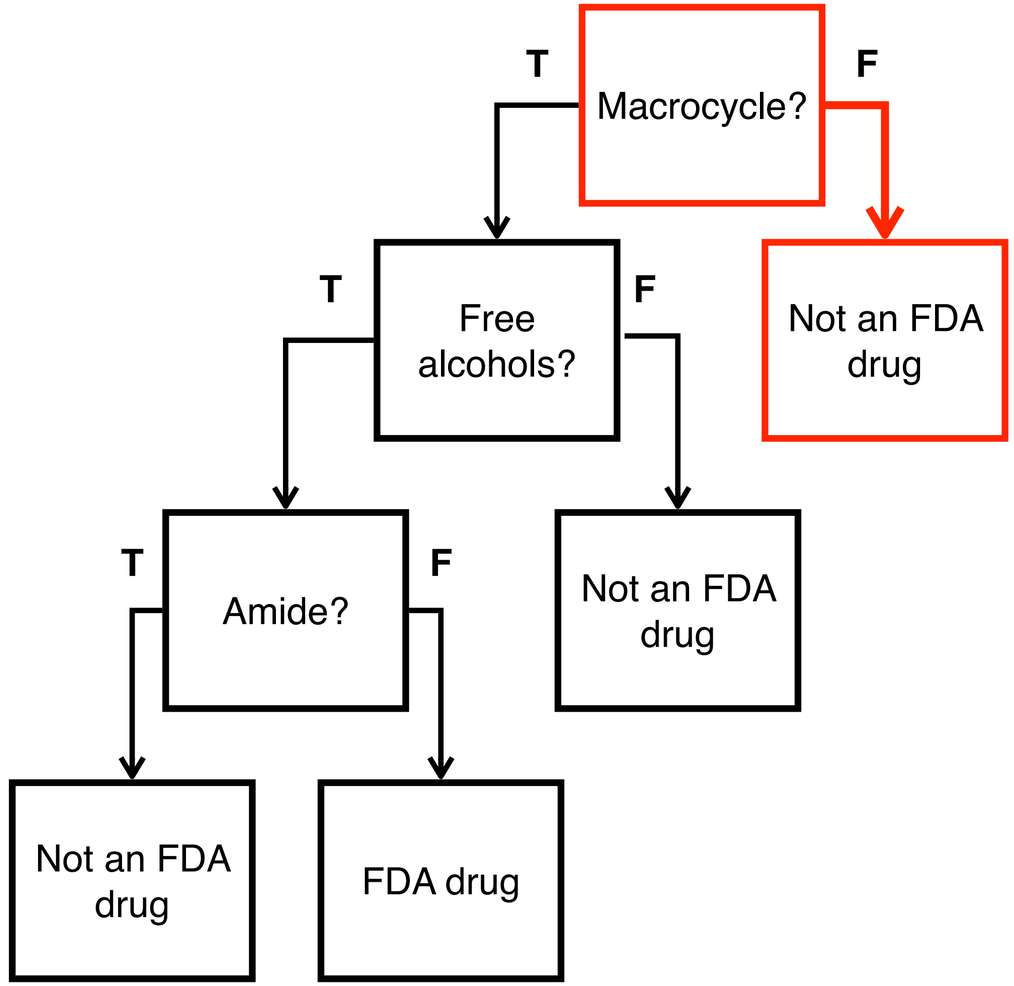
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No

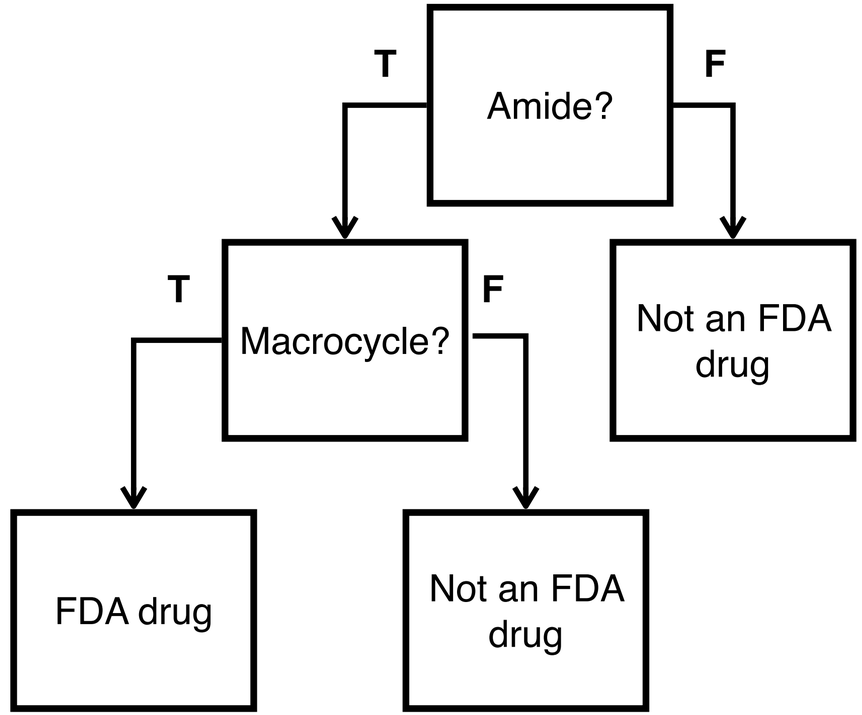


Some decision trees are better than others…

Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

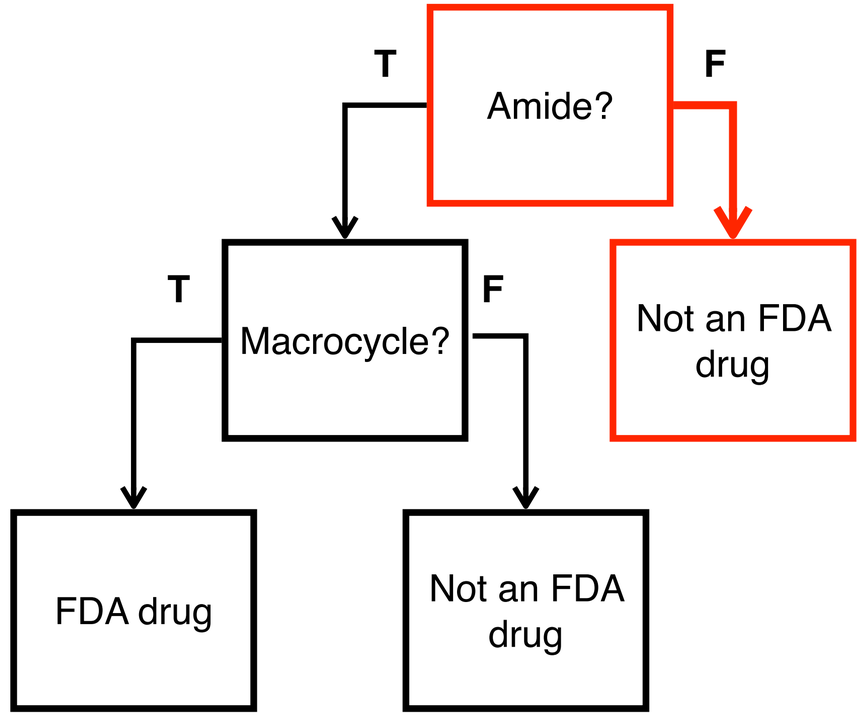
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

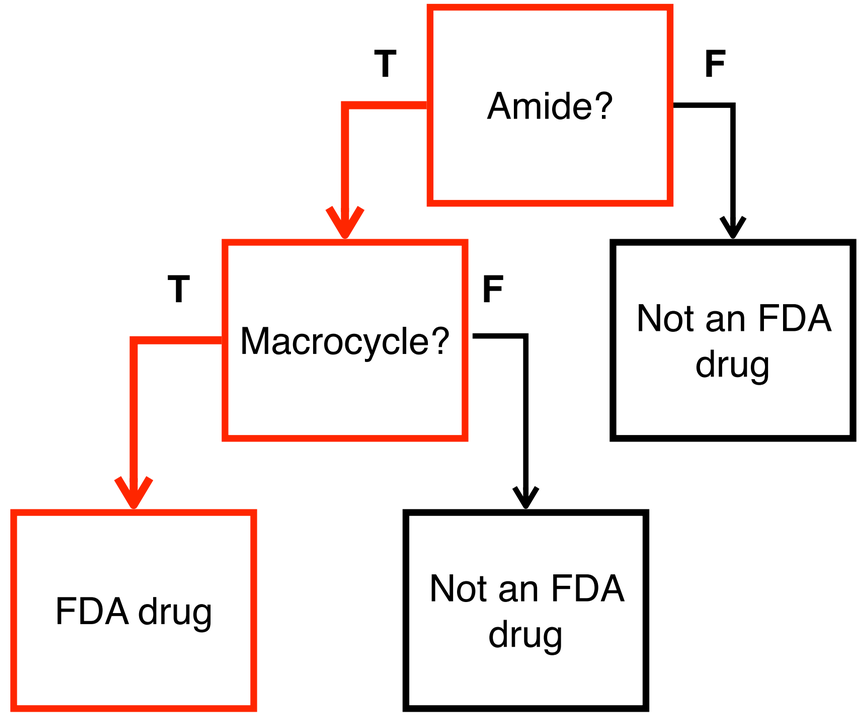
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

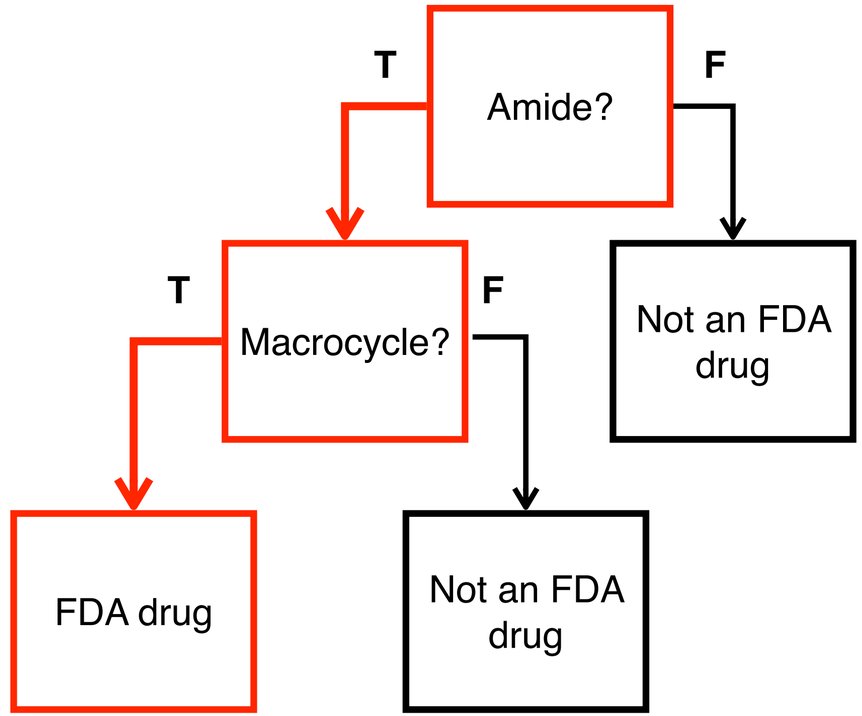
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

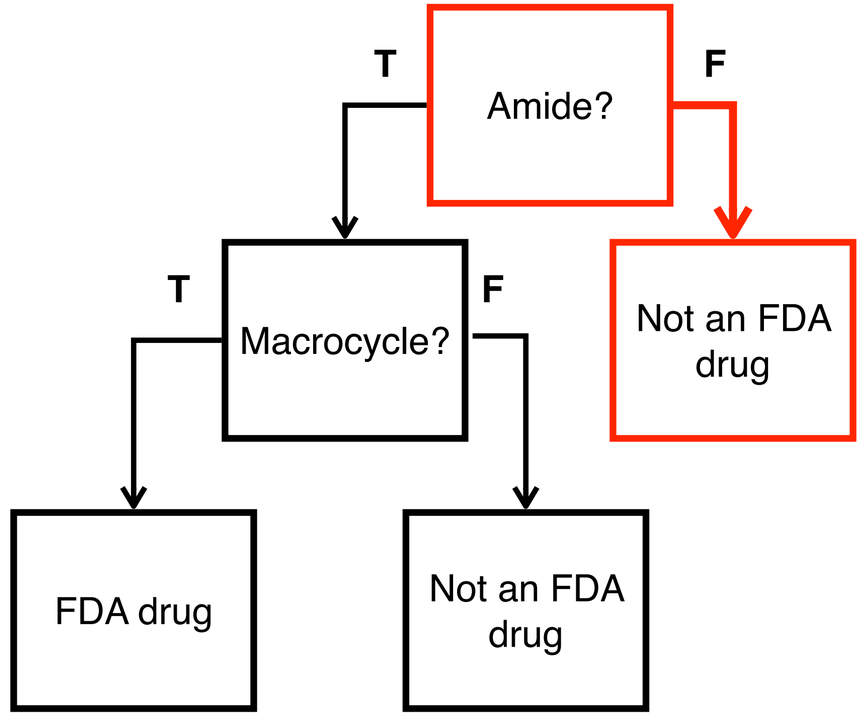
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Classifiers

UNIVERSITY OF CAMBRIDGE

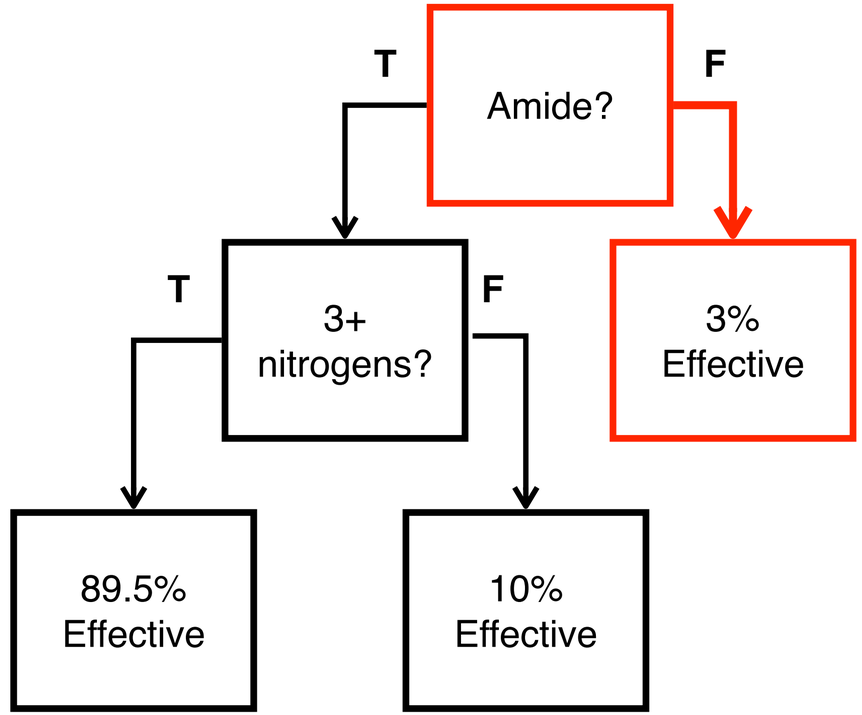
Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision Tree Regressors

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	# Nitrogens	Contains amide?	Antibacterial Effectiveness
1	Yes	2	No	3%
2	No	3	Yes	99%
3	Yes	6	Yes	80%
4	No	1	Yes	10%

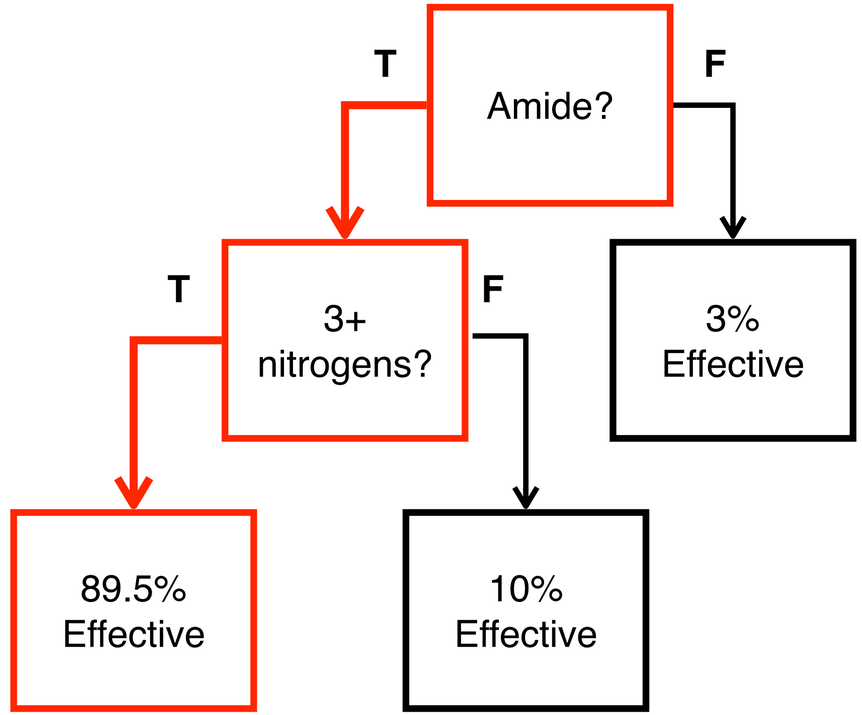


NOTE: This is an oversimplification of regression trees!

Decision Tree Regressors

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	# Nitrogens	Contains amide?	Antibacterial Effectiveness
1	Yes	2	No	3%
2	No	3	Yes	99%
3	Yes	6	Yes	80%
4	No	1	Yes	10%

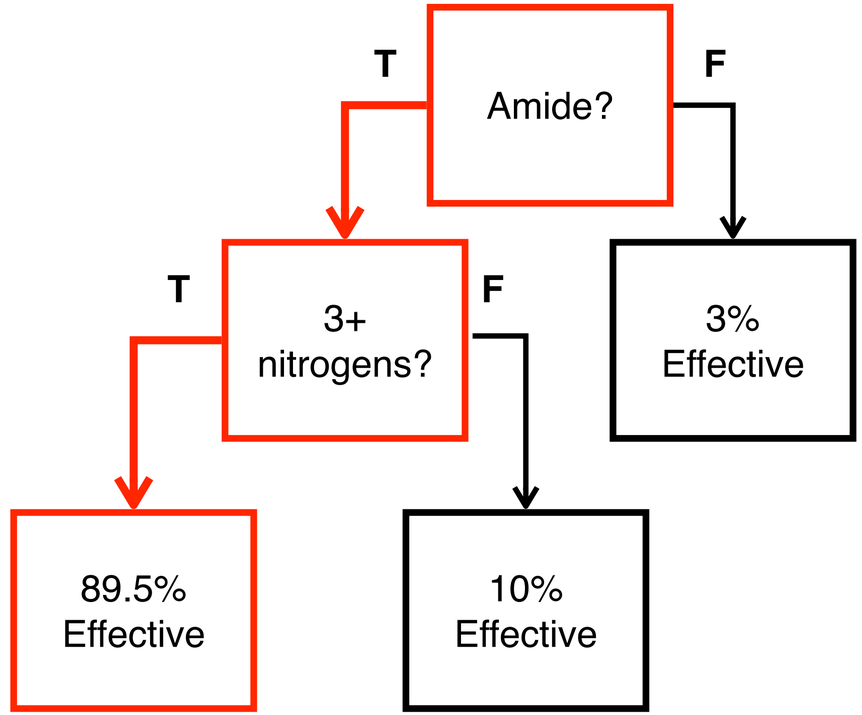


NOTE: This is an oversimplification of regression trees!

Decision Tree Regressors

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	# Nitrogens	Contains amide?	Antibacterial Effectiveness
1	Yes	2	No	3%
2	No	3	Yes	99%
3	Yes	6	Yes	80%
4	No	1	Yes	10%

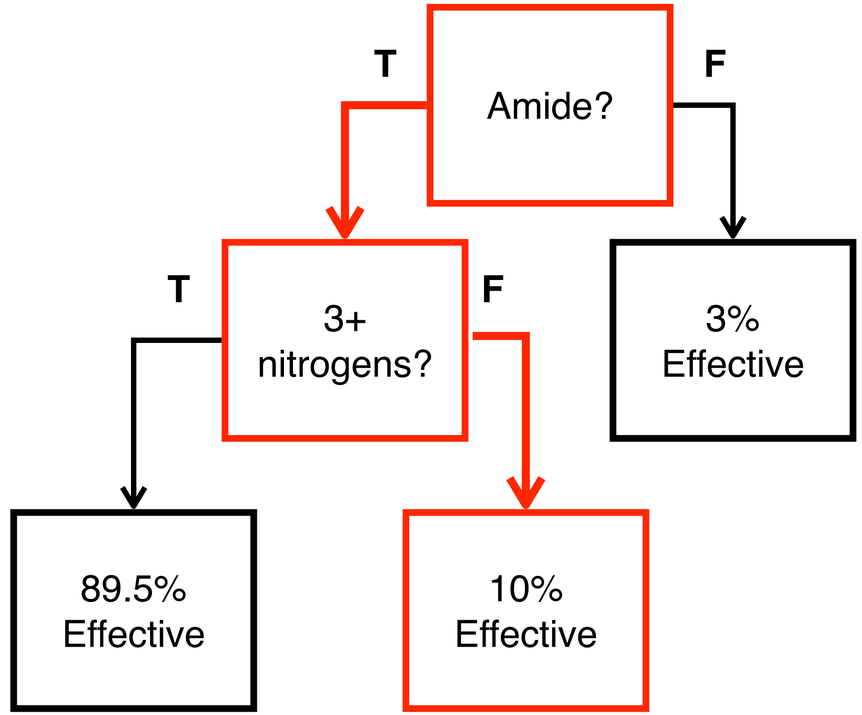


NOTE: This is an oversimplification of regression trees!

Decision Tree Regressors

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	# Nitrogens	Contains amide?	Antibacterial Effectiveness
1	Yes	2	No	3%
2	No	3	Yes	99%
3	Yes	6	Yes	80%
4	No	1	Yes	10%



NOTE: This is an oversimplification of regression trees!

UNIVERSITY OF CAMBRIDGE

Some challenges:

It’s easy to overfit a decision tree - you can just keep adding branches until you perfectly predict everything.

Tend to be highly sensitive to the training data.

Random assortment of decision trees

Use a random subset of the original data

Rely upon the “wisdom of crowds”

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No

Randomly selected rows (can sample the same row more than once)

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?

Randomly selected rows (can sample the same row more than once)

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No

Randomly selected rows (can sample the same row more than once)

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
3	Yes	Yes	Yes	Yes

Randomly selected rows (can sample the same row more than once)

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
3	Yes	Yes	Yes	Yes
1	Yes	No	No	No

Randomly selected rows (can sample the same row more than once)

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
3	Yes	Yes	Yes	Yes
1	Yes	No	No	No
2	No	Yes	Yes	Yes

Randomly selected rows (can sample the same row more than once)

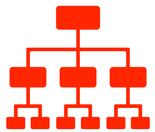
Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No

Bootstrapped Data 1

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
3	Yes	Yes	Yes	Yes
1	Yes	No	No	No
2	No	Yes	Yes	Yes



Decision tree 1

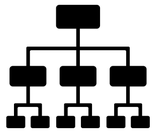
Build a decision tree with **THIS** data.

Improves stability, reduces noise and overfitting

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision tree 1

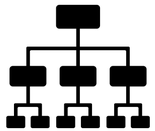
Bootstrapped Data 2

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision tree 1

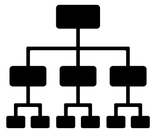
Bootstrapped Data 2

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision tree 1

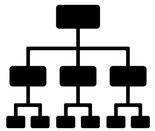
Bootstrapped Data 2

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision tree 1

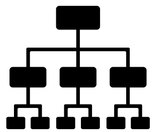
Bootstrapped Data 2

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
2	No	Yes	Yes	Yes

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No



Decision tree 1

Bootstrapped Data 2

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
2	No	Yes	Yes	Yes
2	No	Yes	Yes	Yes

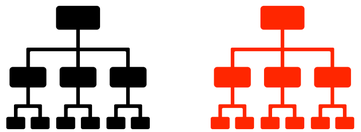
Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No

Bootstrapped Data 2

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
2	No	Yes	Yes	Yes
2	No	Yes	Yes	Yes



Decision tree 1

Decision tree 2

Build a **different** decision tree with **THIS** data.

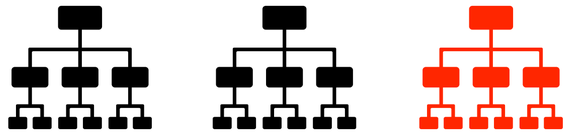
Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No

Bootstrapped Data 3

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
4	No	No	No	No
3	Yes	Yes	Yes	Yes
2	No	Yes	Yes	Yes
2	No	Yes	Yes	Yes



Decision tree 1 | Decision tree 2 | Decision tree 3

Build a **different** decision tree with **THIS** data.

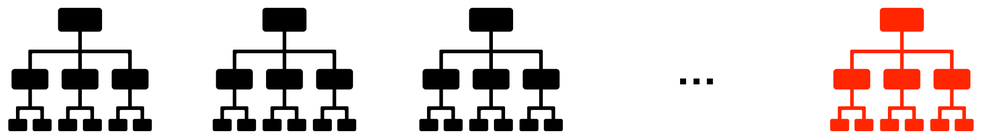
Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes
4	No	No	No	No

Bootstrapped Data N

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
1	Yes	No	No	No
4	No	No	No	No
1	Yes	No	No	No
4	No	No	No	No



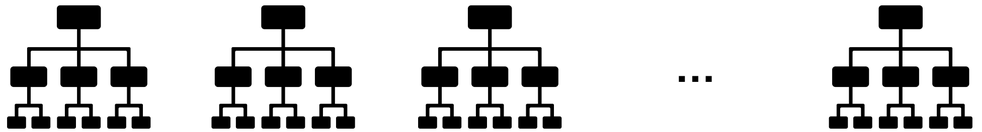
Decision tree 1 | Decision tree 2 | Decision tree 3 | ... | Decision tree N

Build a **different** decision tree with **THIS** data.

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
5	No	Yes	Yes	No

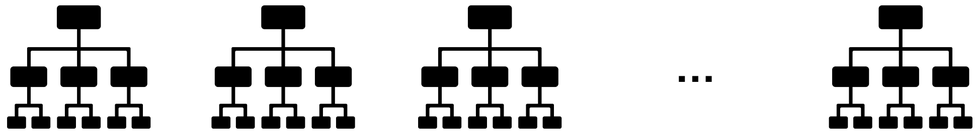


Decision tree 1 | Decision tree 2 | Decision tree 3 | ... | Decision tree N

Random Forests: The Big Idea

UNIVERSITY OF CAMBRIDGE

Molecule	Free alcohols?	Is a Macrocycle?	Contains amide?	Is FDA approved drug?
5	No	Yes	Yes	No



Decision tree 1 | Decision tree 2 | Decision tree 3 | ... | Decision tree N

No | No | Yes | | Yes

If more trees say “No” than “Yes”, molecule is classified as not an FDA approved drug.

Decision Trees vs Random Forests

UNIVERSITY OF CAMBRIDGE

Decision Trees		**Random Forests**
	Classification?	
	Regression?	
	Builds model from entire dataset?	
More prone to overfitting	Overfitting Tendency?	More stable

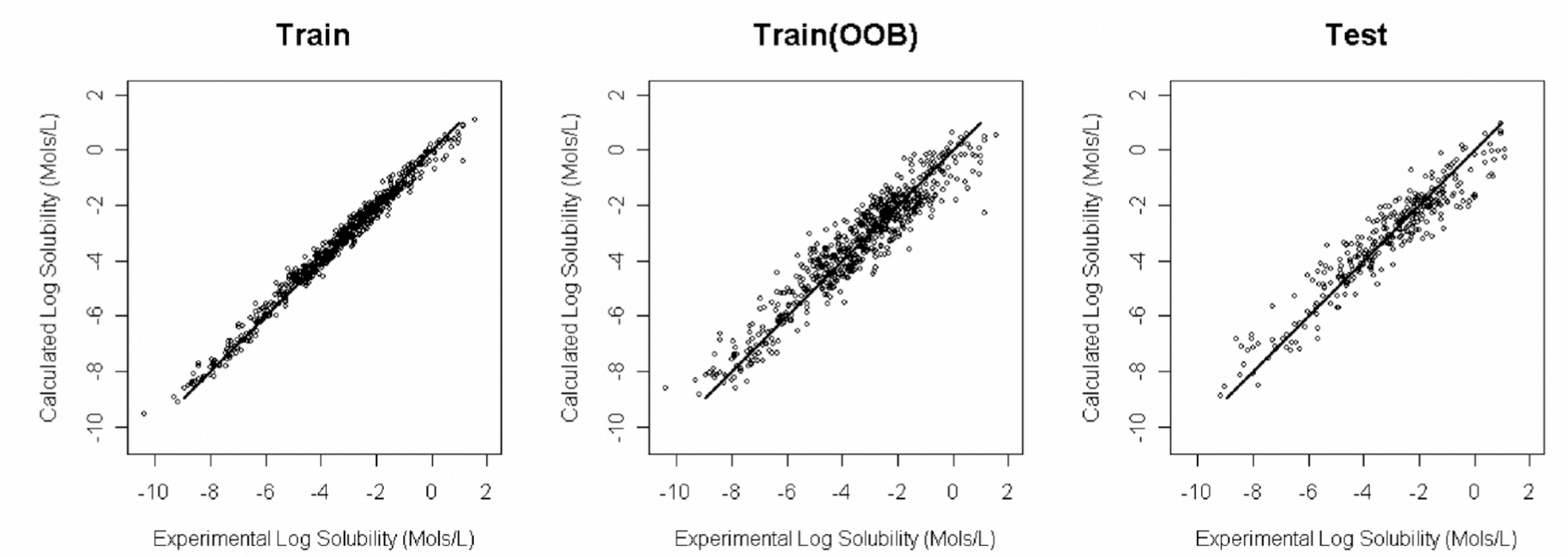
UNIVERSITY OF CAMBRIDGE

ADME Property optimization

Absorption, Distribution, Metabolism, Excretion

Palmer *et al. J. Chem. Inf. Model.* **2007**, *47*, 150.

Some Real World Examples of Random Forests



model	r^2(te)	RMSE(te)
PLS	0.859	0.773
ANN	0.866	0.751
SVM	0.878	0.720
RF	0.890	0.690

Palmer *et al. J. Chem. Inf. Model.* **2007**, *47*, 150.

Some Real World Examples of Random Forests

UNIVERSITY OF CAMBRIDGE

prune

hay

cooked cherry

woody

cherry stones

blackcurrant

leather

bell pepper

musk

black pepper

ethanol

cut grass

fresh cherry

fresh strawberry

blackberry

Some Real World Examples of Random Forests

UNIVERSITY OF CAMBRIDGE

prune

hay

cooked cherry

woody

cherry stones

blackcurrant

leather

bell pepper

musk

black pepper

ethanol

cut grass

fresh cherry

fresh strawberry

blackberry

Some Real World Examples of Random Forests

UNIVERSITY OF CAMBRIDGE

N° CAS	Name	N° CAS	Name	N° CAS	Name
624–41-9	2-methyl-1-butyl acetate	142–92-7	hexyl acetate	60–12-8	phenylethyl alcohol
624–54-4	amyl propanoate	123–92-2	isoamyl acetate	127–41-3	α-ionone
123–86-4	butyl acetate	105–68-0	isoamyl propanoate	23726–93-4	β-damascenone
97–64-3	ethyl 2-hydroxy-propanoate	503–74-2	3-methyl-butanoic acid	79–77-6	β-ionone
7452–79-1	ethyl 2-methyl-butanoate	103–82-2	benzeneacetic acid	2785–89-9	2-methoxy-4-ethylphenol
97–62-1	ethyl 2-methyl-propanoate	107–92-6	butanoic acid	123–07-9	4-ethylphenol
5405–41-4	ethyl 3-hydroxy-butanoate	111–27-3	1-hexanol	431–03-8	2.3-butanedione
108–64-5	ethyl 3-methyl-butanoate	111–87-5	1-octanol	122–78-1	benzeneacetaldehyde
105–54-4	ethyl butanoate	770–35-4	1-phenoxypropan-2-ol	3268–49-3	methional
123–66-0	ethyl hexanoate	137–32-6	2-methyl-butan-1-ol	3658–77-3	furaneol
106–32-1	ethyl octanoate	544–12-7	3-hexen-1-ol	27538–10-9	homofuraneol
105–37-3	ethyl propanoate	123–51-3	3-methyl-butanol	24683–00-9	IBMP
5299–60-5	ethyl-6-hydroxyhexanoate	505–10-2	methionol		

Vigneau *et al. Food. Qual. Prefer.* **2018**, *68*, 135.

Some Real World Examples of Random Forests

UNIVERSITY OF CAMBRIDGE

	Bell Pepper			Leather		
	#VOCs	#comp.	RMSEcv	#VOCs	#comp.	RMSEcv
Null model	–	–	0.412	–	–	0.387
RF	38	–	0.362	38	–	0.317
RF with selected VOCs	3	–	0.259	2	–	0.263
PLS Reg.	38	1	0.366	38	1	0.338
PLS Reg. (V.I.P. > 1.2)	8	7	0.295	14	2	0.289

Vigneau *et al. Food. Qual. Prefer.* **2018**, *68*, 135.

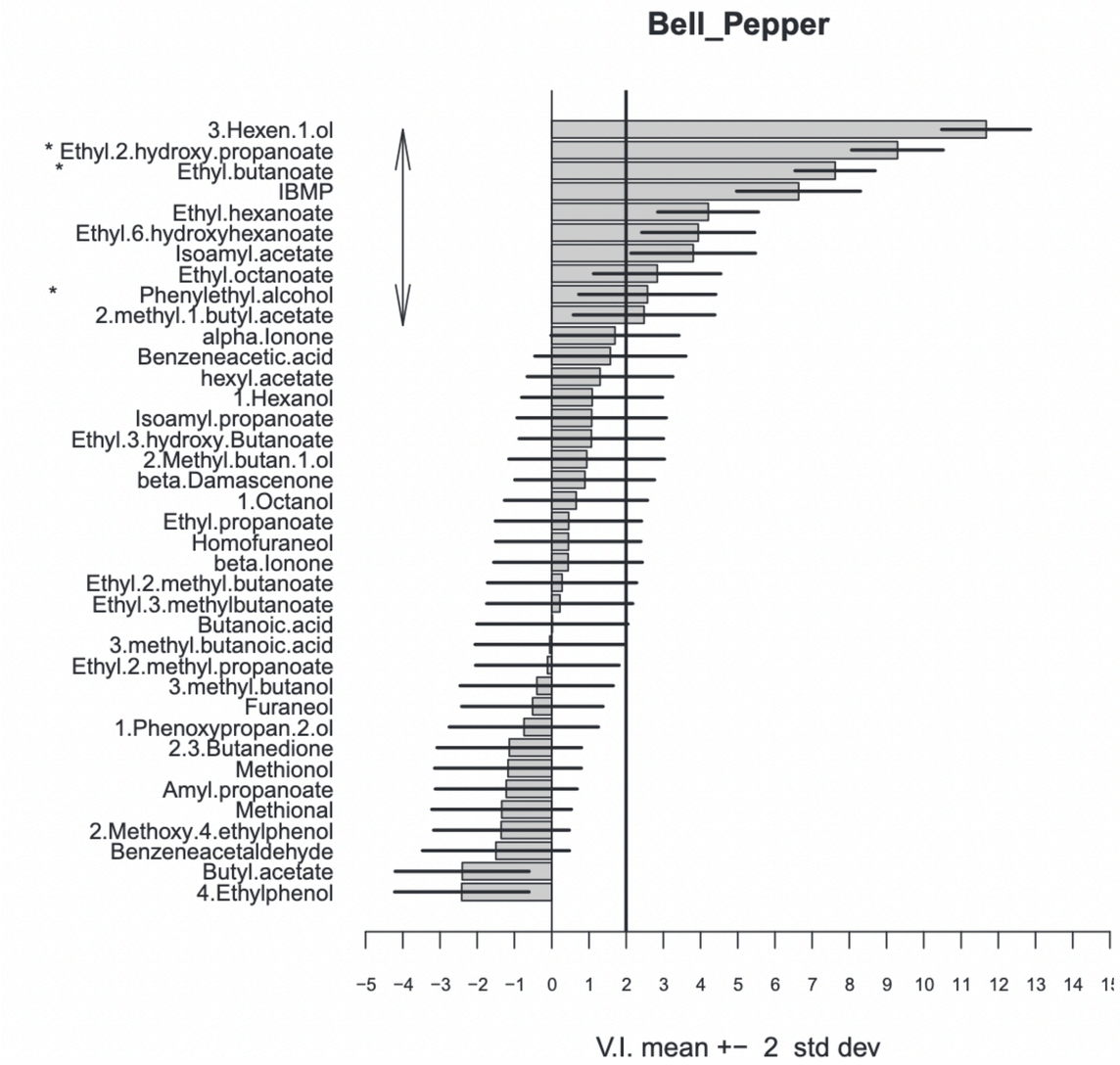
Some Real World Examples of Random Forests

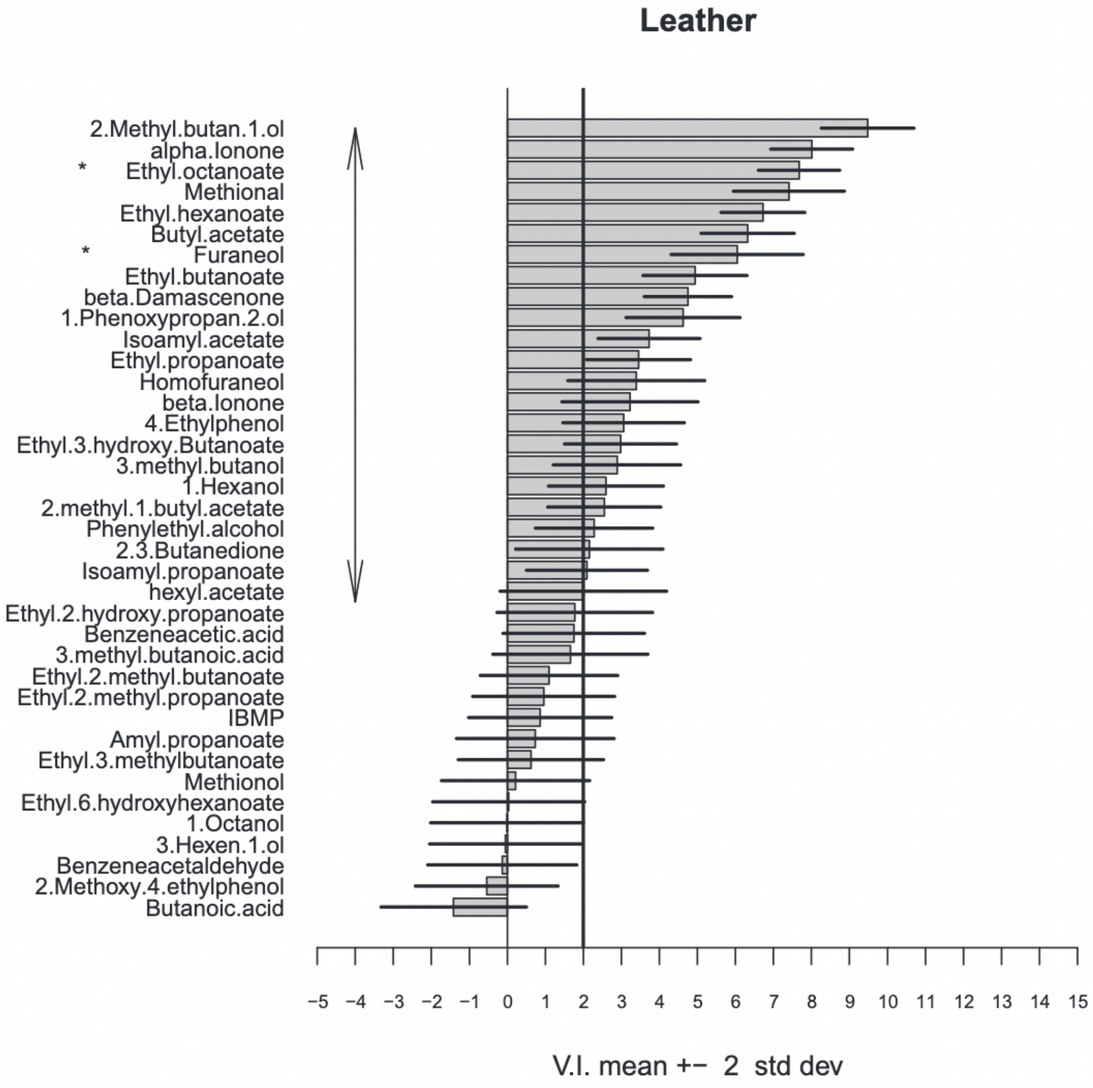
UNIVERSITY OF CAMBRIDGE

	Bell Pepper			Leather		
	#VOCs	#comp.	RMSEcv	#VOCs	#comp.	RMSEcv
Null model	–	–	0.412	–	–	0.387
RF	38	–	0.362	38	–	0.317
RF with selected VOCs	3	–	0.259	2	–	0.263
PLS Reg.	38	1	0.366	38	1	0.338
PLS Reg. (V.I.P. > 1.2)	8	7	0.295	14	2	0.289

Vigneau *et al. Food. Qual. Prefer.* **2018**, *68*, 135.

Some Real World Examples of Random Forests





Vigneau *et al. Food. Qual. Prefer.* **2018**, *68*, 135.