

Decision Trees

You are stranded on a deserted island. Mushrooms of various types grow widely all over the island, but no other food is anywhere to be found. Some of the mushrooms have been determined as poisonous and others as not (determined by your former companions' trial and error!). You are the only one remaining on the island. You have the data in the Table to consider.

1. Build a Decision Tree to classify if a mushroom is edible or not.
2. You know whether mushrooms from *A* to *H* are edible or not, but you need to figure out about *X*, *Y* and *Z* type. Use the above built DT to classify them.
3. Would ever eat any of *X*, *Y*, *Z*? Why?

Example	IsHeavy	IsSmelly	IsSpotted	IsSmooth	IsEdible
<i>A</i>	1	0	0	0	1
<i>B</i>	1	0	1	0	1
<i>C</i>	0	1	0	1	1
<i>D</i>	0	0	0	1	0
<i>E</i>	1	1	1	0	0
<i>F</i>	1	0	1	1	0
<i>G</i>	1	0	0	1	0
<i>H</i>	0	1	0	0	0
<i>X</i>	0	1	1	1	?
<i>Y</i>	1	1	0	1	?
<i>Z</i>	1	1	0	0	?

Rocchio classification

Given a series of Documents, described by a text and an assigned class:

1. Find profiles for the class C1 and C2.
2. Find the decision boundaries between the two classes that are consistent with the training data
3. Estimate values for β and γ . that realize property 2.
4. Classify D8.
5. For which values of ω_1 and ω_2 D9 is in class C1?

In the Table, the text is a series of symbols, here simplified as *As* and *Bs*.

D8 is described by 14 *As* and 36 *Bs*.

Document	Text	Class
D1	A A B B B B B B B	C1
D2	A A A B B B B B B B B	C1
D3	A B B B B B B	C1
D4	A A A A A A B B B B	C2
D5	A A A A A A A A B B B B B B	C2
D6	A A A A A B B	C2
D7	A A A A A A A A A B B B B	C2
D8	14A, 36B	?
D9	$\omega_1 A, \omega_2 B$	C1