

Classification

rule 1

~~declare~~
declare
spam
iff

$$(\hat{p}_i > 0.5)$$

STAT 206
20 Mar 21
extra
DD office
how

for email i

(2)

declare
spam
if

entire ~~95%~~ ^{99.9%}
interval for
 \hat{p}_i is > 0.5

(1)

truth

no sp

spam not spam

	spam	not spam
spam	✓	n_{FP}
not spam	n_{FN}	✓

$$\text{'error'} = \frac{n_{FP} + n_{FN}}{n_V}$$

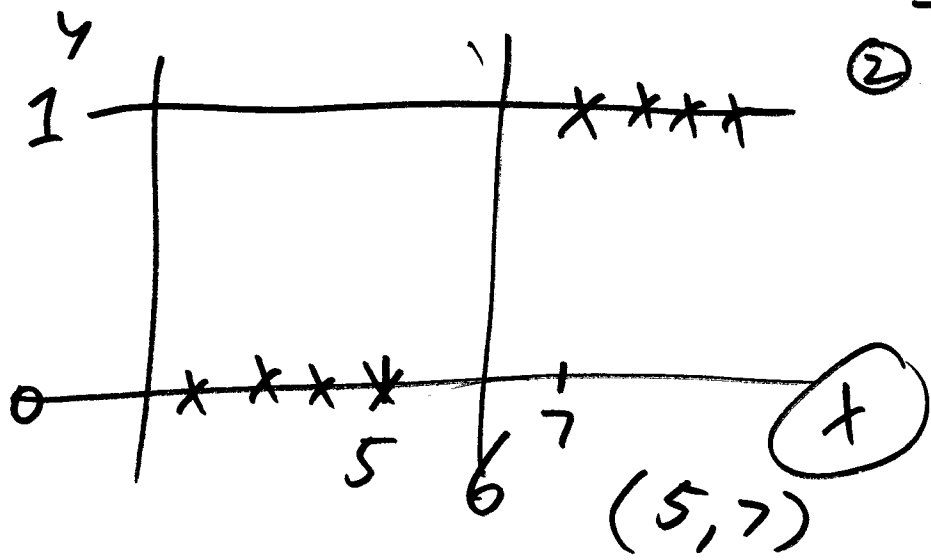
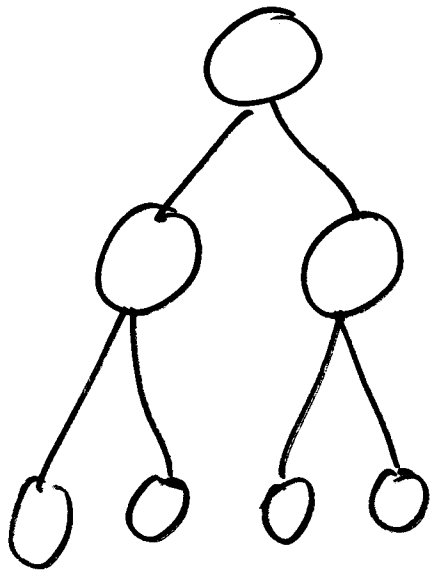
$$1536 = n_V$$

CART : Classification & Regression
 (1985) Y { ! } Trees

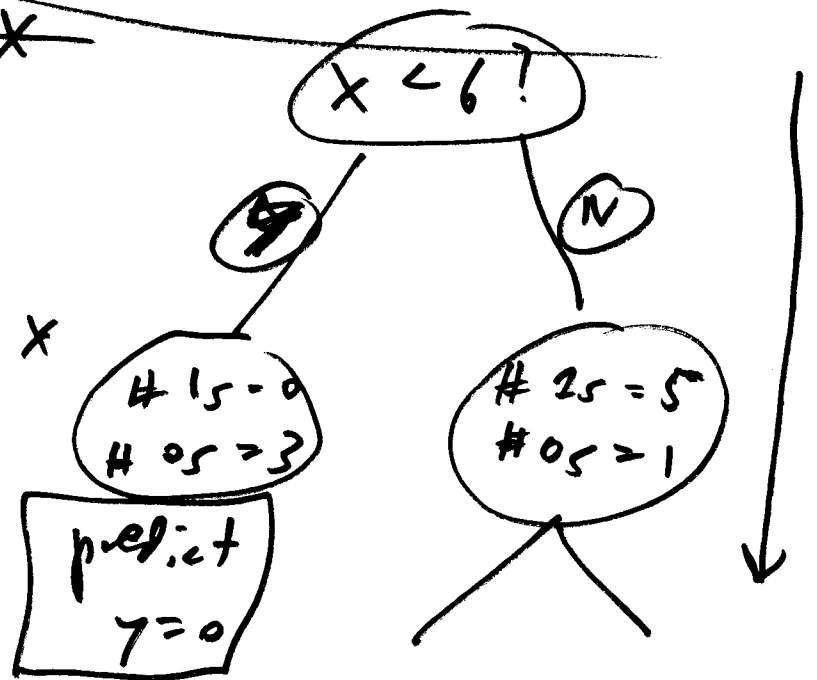
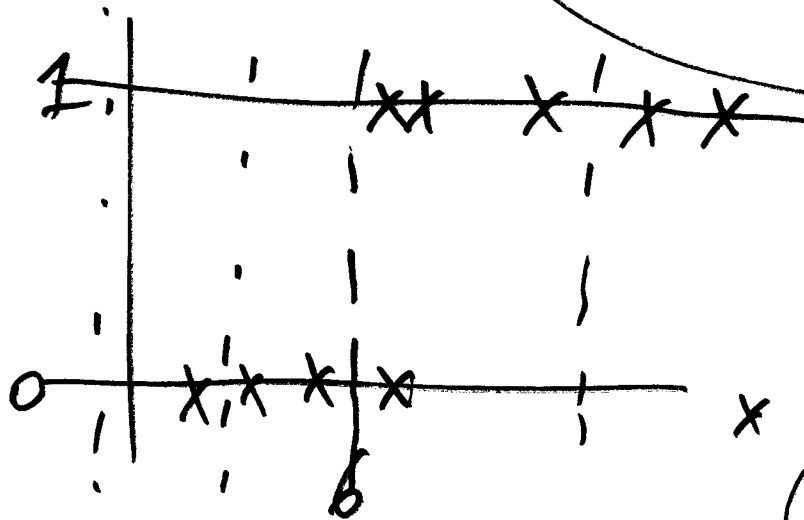
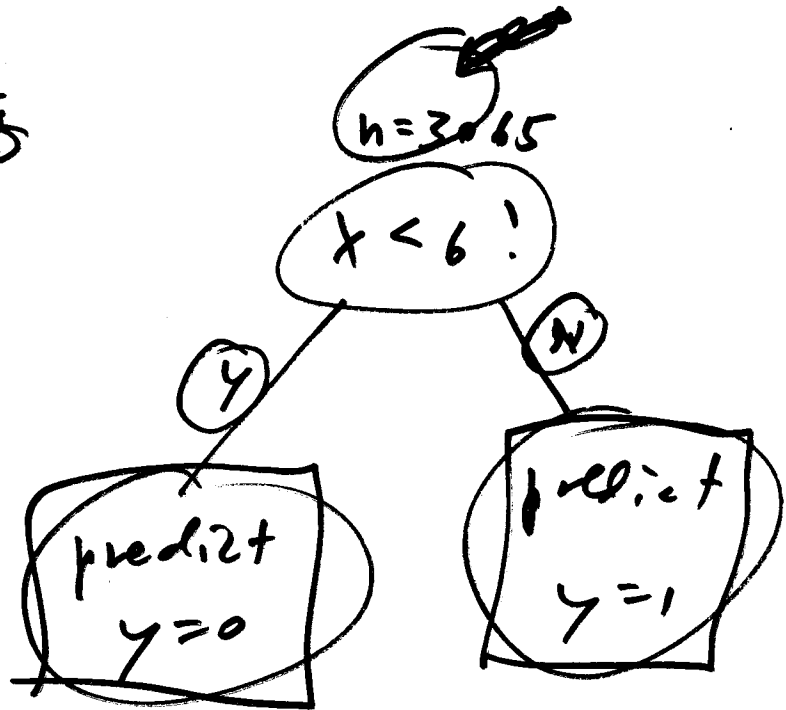
$$y = \begin{cases} 1 \\ 0 \end{cases}$$

binary trees

x



✂



'odor' only

(.05) truth

all say

poisonous edible

(p)	3796 TP	0 FP
(e)	1228 FN	4208 TN

bad (people avoid eating good mushroom)

to achieve 0

really bad

(people encouraged to eat (1) mushroom)

in FN, must suffer lots of FP

'odor' only (.01) truth

(p)	3916	3408
(e)	0	500

total 530