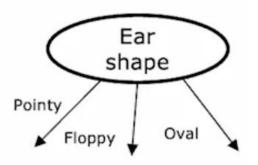
# DECISION TREE LEARNING



# O Features with three possible values

_	Ear shape $(x_1)$	Face shape $(x_2)$	Whiskers $(x_3)$	Cat (y)
	Pointy	Round	Present	1
	Oval	Not round	Present	1
<b>3</b>	Oval	Round	Absent	0
	Pointy	Not round	Present	0
(3)	Oval	Round	Present	1
<b>(2)</b>	Pointy	Round	Absent	1
<b>3</b>	Floppy	Not round	Absent	0
	Oval	Round	Absent	1
VEY!	Floppy	Round	Absent	0
	Floppy	Round	Absent	0



# One hot Encoding

	Ear shape	Pointy ears	Floppy ears	Oval ears	Face shape	Whiskers	Cat
2.07	Pointy				Round	Present	1
No.	Oval				Not round	Present	1
3	Oval				Round	Absent	0
	Pointy				Not round	Present	0
3	Oval				Round	Present	1
	Pointy				Round	Absent	1
:3	Floppy				Not round	Absent	0
3	Oval				Round	Absent	1
3	Floppy	***************************************	***************************************		Round	Absent	0
	Floppy				Round	Absent	0

# One hot Encoding

Ear	shape	Pointy ears	Floppy ears	Oval ears	Face shape	Whiskers	Cat
- P	ointy	1	0	0	Round	Present	1
E Company	<del>Oval</del>	0	0	1	Not round	Present	1
	<del>Oval</del>	0	0	1	Round	Absent	0
P	ointy	1	0	0	Not round	Present	0
Va	<del>Oval</del>	0	0	1	Round	Present	1
P	<del>ointy</del>	1	0	0	Round	Absent	1
70	<del>loppy</del>	0	1	0	Not round	Absent	0
3	<del>Oval</del>	0	0	1	Round	Absent	1
F	<del>loppy</del>	0	1	0	Round	Absent	0
F F	<del>loppy</del>	0	1	0	Round	Absent	0



#### One hot Encoding

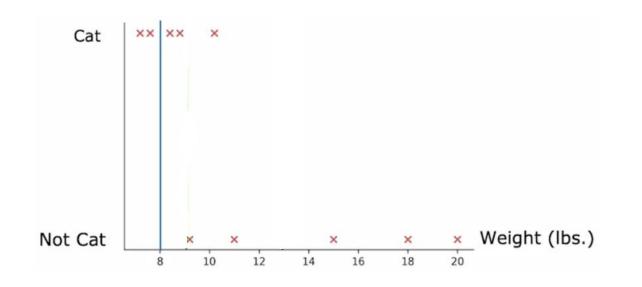
• If a categorical feature can take on k values, create k binary features (0 or 1 valued).

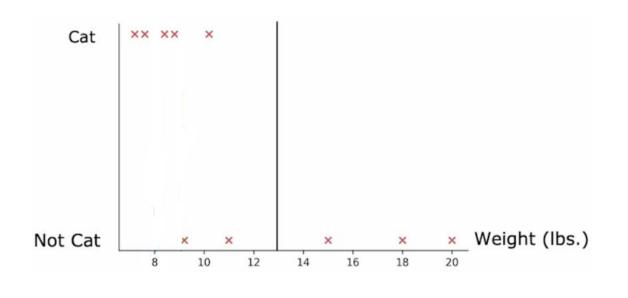


#### Continues features

	Ear shape	Face shape	Whiskers	Weight (lbs.)	Cat
3	Pointy	Round	Present	7.2	1
	Floppy	Not round	Present	8.8	1
(F)	Floppy	Round	Absent	15	0
:	Pointy	Not round	Present	9.2	0
(E)	Pointy	Round	Present	8.4	1
	Pointy	Round	Absent	7.6	1
(E)	Floppy	Not round	Absent	11	0
(=)	Pointy	Round	Absent	10.2	1
( )	Floppy	Round	Absent	18	0
	Floppy	Round	Absent	20	0

## Splitting on a continues variable

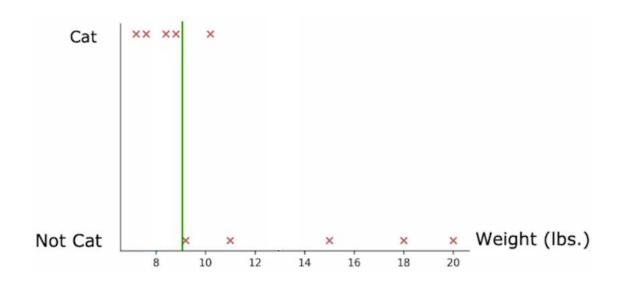


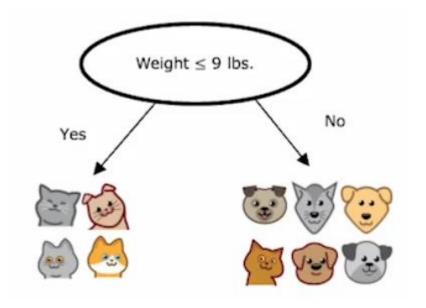


$$H(0.5) - \left(\frac{2}{10}H\left(\frac{2}{2}\right) + \frac{8}{10}H\left(\frac{3}{8}\right)\right) = 0.24$$

$$H(0.5) - \left(\frac{7}{10}H\left(\frac{5}{7}\right) + \frac{3}{10}H\left(\frac{0}{3}\right)\right) = 0.40$$

## O Splitting on a continues variable





$$H(0.5) - \left(\frac{4}{10}H\left(\frac{4}{4}\right) + \frac{6}{10}H\left(\frac{1}{6}\right)\right) = 0.61$$

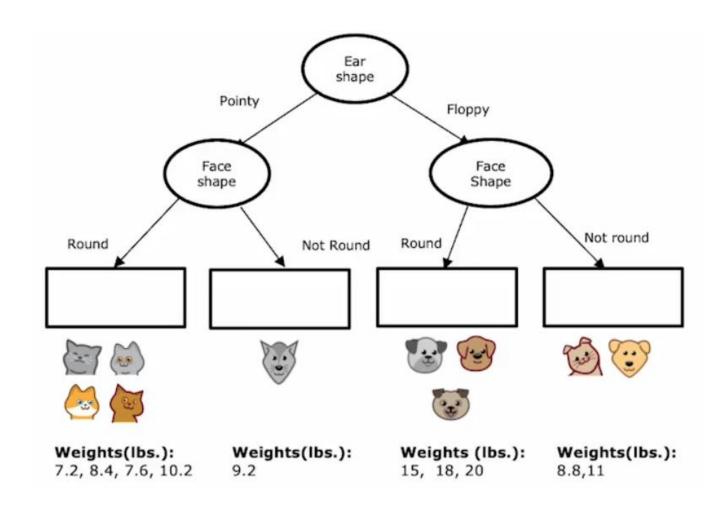


# Regression with Decision Trees: Predicting a number

2	Ear shape	Face shape	Whiskers	Weight (lbs.)
(E)	Pointy	Round	Present	7.2
	Floppy	Not round	Present	8.8
3	Floppy	Round	Absent	15
8.3	Pointy	Not round	Present	9.2
(E)	Pointy	Round	Present	8.4
	Pointy	Round	Absent	7.6
(E)	Floppy	Not round	Absent	11
(3)	Pointy	Round	Absent	10.2
( )	Floppy	Round	Absent	18
	Floppy	Round	Absent	20

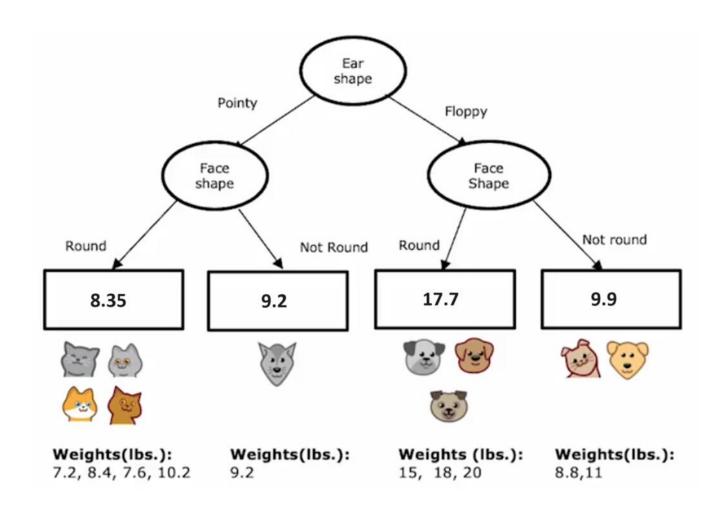


#### Regression with Decision Trees



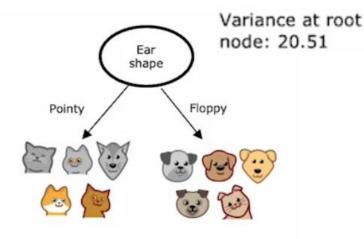


#### Regression with Decision Trees





#### Choosing a Split



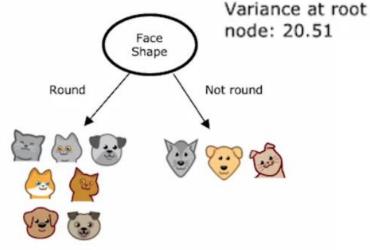
Weights: 7.2, Weights: 8.8, 15, 9.2, 8.4, 7.6, 10.2 11, 18, 20

Weights: 7.2, Weights: 8.8, 15, 9.2, 8.4,7.6, 10.2 11, 18, 20

Variance: 1.47 Variance: 21.87

$$w^{\text{left}} = \frac{5}{10}$$
  $w^{\text{right}} = \frac{5}{10}$ 

$$20.51 - \left(\frac{5}{10} * 1.47 + \frac{5}{10} * 21.87\right)$$
$$= 8.84$$



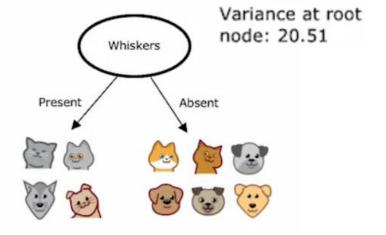
Weights: 7.2, 15, Weights: 8.8,9.2,11 8.4, 7.6,10.2, 18, 20

Weights: 7.2, 15, Weights: 8.8,9.2,11 8.4, 7.6,10.2, 18, 20

Variance: 27.80 Variance: 1.37

 $w^{\text{left}} = \frac{7}{10}$   $w^{\text{right}} = \frac{3}{10}$ 

$$20.51 - \left(\frac{7}{10} * 27.80 + \frac{3}{10} * 1.37\right)$$
$$= 0.64$$



Weights: 7.2, 8.8, Weights: 15, 7.6, 9.2, 8.4 11, 10.2, 18, 20

Weights: 7.2, 8.8, Weights: 15, 7.6, 9.2, 8.4 11, 10.2, 18, 20

Variance: 0.75 Variance: 23.32

 $w^{\text{left}} = \frac{4}{10}$   $w^{\text{right}} = \frac{6}{10}$ 

$$20.51 - \left(\frac{4}{10} * 0.75 + \frac{6}{10} * 23.32\right)$$

$$= 6.22$$