

Decision Tree

(Row 2, Row 3, Row 4, Row 5) \times Train

(Row 6) \times Train

$$(1) \text{ Values} = 1725 + 4449 + 21984 + 3866$$

(2 - 5)

$$\text{Average} = 8006$$

$$\begin{matrix} \text{Starting} \\ \text{Messiness} \end{matrix} \quad (\text{variance}) = 66, 155, 786$$

(2) Finding Best Question
Option A: split by Age (If Age < 30)

Left room : Row 2 (18), Row 3 (28)
(Young)
Cost : 1725, 4449

Messiness scores:

Right Room : Row 4 (33), Row 5 (32)
Cost : 21984, 3, 866
(Room average) 4 | 1958 459 (verbal)
Messiness score 1

Option B (split by BMI) (< 25)

Left room: 0

Room Average: 1371588 (very low) (win)

Right room:

(3) Building Leaf nodes

Bucket (1)

(skinny and expensive gp)

Prediction: 21984

Bucket (2)

(the high BMI gp)

Prediction: $\frac{1725 + 4449 + 3866}{3}$
: 3347

(4) Test Row 6

Age: 31

BMI: 25.74

Y test Actual Cost: \$3756

Predicted Path

1. Is your BMI < 25
2. go right
3. \$3347 (model)

* off only by \$409

Messiness Score is get?

1. Find mean : $\frac{1726 + 4449 + 21984 + 3867}{4}$

; .8006 .59

2. mean Squared error for each person

Person(1) = $(1726 - 8007)^2 = 39451372$

Person(2) = $(4449 - 8007)^2 = 17137363$

Person(3) = $(21984 - 8007)^2 = \text{result}$
mean Squared error = $\frac{\text{result}}{4 \text{ people}}$

3. Average mean Squared error = $= 66155786$

* High variance = Bad

* Low variance = Good

Messiness score

How to find winner

Left room, (1 out of 4 people) = 25.1. , 0

Right room, (3 out of 4 people) = 75.1. 1371588

Right room,

$$\text{Total Score} = (0.25 \times 0) + (0.75 \times 137/588)$$
$$= 1028.691$$

* starting score: 66155786

* Age split score: 41958459

* BMI split score: 1028691 ✓

