1).

imp

	Coefficient	Std. error	t-statistic	p-value
Intercept	2.939	0.3119	9.42	< 0.0001
TV	0.046	0.0014	32.81	< 0.0001
radio	0.189	0.0086	21.89	< 0.0001
newspaper	−0.001	0.0059	−0.18	0.8599

→ Large

TABLE 3.4. *For the* Advertising *data, least squares coefficient estimates of the multiple linear regression of number of units sold on TV, radio, and newspaper advertising budgets.*

NULL HYPOTHESIS :-

↓

There is no R^n

Simple Regression $\begin{cases} H_0 : \beta_1 = 0 \\ H_a : \beta_1 \neq 0 \end{cases}$

Mutiple Regression $\begin{cases} H_0 : \beta_1 = \beta_2 \dots \beta_p = 0 \\ H_a : \text{atleast one } \beta_j \text{ is non zero} \end{cases}$

A : Null Hypo :

- TV has no R^n wrt Sales
- Rad has no R^n wrt sales
- News has no R^n wrt sales

A: conclusion :

- As the p value for TV & radio is very less. we can say that there is R^n between these parameters and Sales.
- P (Newspaper) ...

... is right $\therefore$ do not reject H_0

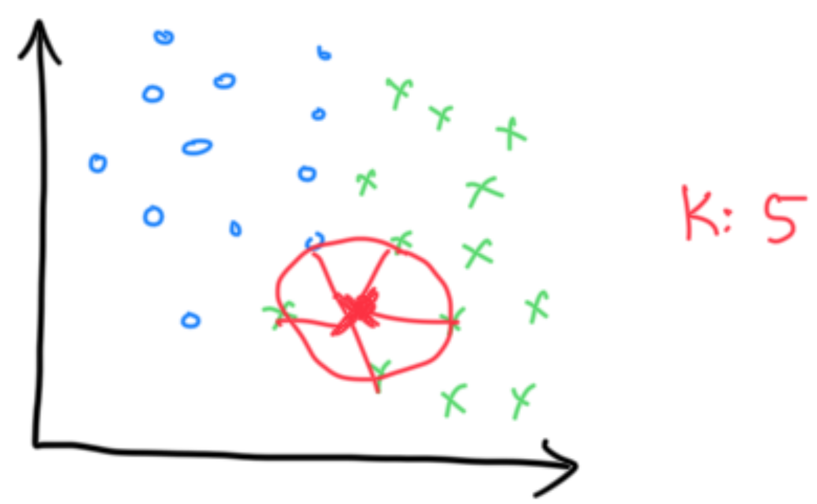
---

2). KNN classifier:

1. get the data
2. Init the K
3. Find nearest K point to our target point

For a class j and point i:

$$P(Y_j \mid X_i) = \frac{1}{K} \sum_{i = N_0} I(y_i = j)$$



$$\therefore P(Y_{blue} \mid X) = \frac{1}{5}(1)$$

$$P(Y_{green} \mid X) = \frac{1}{5}(4)$$

KNN regression:-

Same:

$$\hat{f}(x_0) = \frac{1}{K} \sum_{x_i \in N_0} y_i$$

Find dist btw points

mean - L1 Loss

median - L2 loss

---

3). X_1 : GPA X_2 : IQ X_3 : $\begin{cases} \text{college} : 1 \\ \text{School} : 0 \end{cases}$ X_4 : X_1 & X_2 X_5 = X_1 & X_3

$\beta_0 : 50$ $\beta_1 : 20$ $\beta_2 : 0.07$ $\beta_3 : 35$ $\beta_4 : 0.01$ $\beta_5 : -10.$

Y(starting salary) in 1000 $

A :

$$Salary_{coll} : \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \ldots \longrightarrow (1)$$

$$: 50 + 20X_1 + 0.07X_2 + 35X_3 + 0.01X_4 + (-10X_5)$$

$$Salary_{HS} : 50 + 20X_1 + 0.07X_2 + 0 + 0.01X_4 + 0$$

$$Salary_{coll} = Salary_{HS}$$

$$\cancel{50} + 20\cancel{X_1} + 0.07\cancel{X_2} + 35X_3 + 0.01\cancel{X_4} - 10X_5 = \cancel{50} + \cancel{20X_1} + 0.\cancel{07}X_2 + 0.\cancel{01X_4}$$

$$35X_3 - 10X_5 = 0$$

$$35X_3 > 10X_5$$

$$35\cancel{X_3} > 10X_1\cancel{X_3}$$

$$\frac{35}{10} < X_1 \qquad GPA > 3.5 //$$

Only (iii) is correct.

(b) IQ = 110 GPA = 4.0

Plug these into (1)

$ 137100.0

(c) The coeff tells that how much effect it has on unit increase.

Every unit $\uparrow$ of β_4, y $\uparrow$ 10 $

(1) Do S.E to check how good our estimate is

S.E $\lll$ $\rightarrow$ reliable params

$\beta_p \lll$ & S.E $\lll$ = affects $\hat{y}$

$\beta_p \lll$ & S.E $\ggg$: not affecting $\hat{y}$

---

(4). $n = 100$

Linear : $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$

Cubic : $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2^2 \ldots + \epsilon$

(a) X & Y is linear

$RSS_{train\,(lin)}$ $RSS_{train\,(cubic)}$ what would be it ?

From my learning :

RSS_{cubic} will be less than linear because

$\boxed{RSS = e_1^2 + e_2^2 + e_3^2}$

Bias : error of Train data

Var : error of test data

While training linear would have High B & High Var.

training cubic would have low bias & High var.

$RSS_{train\,linear} > RSS_{train\,cubic}$

(b) For testing :

cubic - overfits & linear . underfits

RSS (linear) < RSS (cubic)

It explains more variance

(c) RSS training $\downarrow$ as the model is of high variance

RSS cubic <<< New model

(d) For test we cannot say because If the new model is little more poly the variance $\uparrow$ It