7

SP- LIME

$c(w, V, I) \rightarrow$ Cost function associated with V

$$c(w, V, I) = MSE - \lambda |V|$$

where $MSE = \frac{1}{n} \sum (I(w) - f(w))^2$

I - Complex Model.

V - Set of features used in Explanation.

W - Instance that need to be explained.

$V \leq \underline{\underline{B}} \rightarrow$ maximum No of features to be included in Explanation.

λ - regularization parameter.

$I(w)$ (complex Model) - logistic Regression Model.

$f(w)$ - linear Regression

$$I(w) = \frac{1}{1 + e^{-(w_1 x_1 + w_2 x_2 + w_3 x_3 + b)}}$$

$$f(w) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots \dots \beta_n x_n.$$

Given

$\beta_0 = 10$ $\beta_1 = 12$ $\beta_2 = 15$ $\beta_3 = 11.5$

$w_1 = 0.5$ $w_2 = 0.2$ $w_3 = 0.1$ $b = -2$

I/P $\rightarrow$ 29, 9.4, 22

I Generate perturbed Data Points.

	f_1	f_2	f_3
x_1	27	7.3	21
x_2	28	7.4	23
x_3	32	9.2	31
x_4	15	4.5	15
x_5	33	11.4	21.

II Calculate Complex Model Prediction $I(w_1)$

$I(\omega)$ (for all three features)

$$x_1 = \frac{1}{1+e^{-(0.5(27)+0.2(7.3)+0.1(21)-2}}$$

$$= \frac{1}{1+e^{-(13.5+1.46+2.1-2)}}$$

$$= \frac{1}{1+e^{-15.06}} = 0.9999$$

$$x_2 = \frac{1}{1+e^{-(0.5(28)+0.2(7.4)+0.1(23)-2}}$$

$$= \frac{1}{1+e^{-15.78}} = 0.9999$$

$$x_3 = \frac{1}{1+e^{-(0.5(32)+0.2(9.2)+0.1(31)-2}}$$

$$= \frac{1}{1+e^{-18.94}} = 0.9999$$

$$x_4 = \frac{1}{1+e^{-(0.5(15)+0.2(4.5)+0.1(15))}}$$

$$= \frac{1}{1+e^{-7.9}} = 0.9999$$

$$x_5 = \frac{1}{1+e^{-(0.5(33)+0.2(11.4)+0.1(21)-2}}$$

$$= \frac{1}{1+e^{-18.88}}$$

$$= 0.9999$$

III Compute $f(w_i) \rightarrow$ linear Regression Model Prediction.

$$f(w_i) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

$$x_1 = 10 + 12(27) + 15(7.3) + 11.5(21)$$

$$= 10 + 324 + 109.5 + 241.5$$

$$= 685$$

$$x_2 = 10 + 12(28) + 15(7.4) + 11.5(23)$$

$$= 10 + 336 + 111 + 264.5$$

$$= 721.5$$

$$x_3 = 10 + 12(32) + 15(9.2) + 11.5(31)$$

$$= 10 + 384 + 138 + 356.5 = 888.5$$

$$x_4 = 10 + 12(15) + 15(4.5) + 11.5(15)$$

$$= 10 + 180 + 67.5 + 172.5 = 430$$

$$x_5 = 10 + 12(33) + 15(11.4) + 11.5(21)$$

$$= 10 + 396 + 171 + 241.5 = 818.5$$

$I(w_i)$	$f(w_i)$
0.999	685
0.999	721.5
0.999	888.5
0.999	430
0.999	818.5

IV MSE Computation.

$$MSE = \frac{1}{n} \Sigma \left(I(\omega) - f(\omega) \right)^2$$

$$= \frac{1}{5} \Big[(0.99 - 685)^2 + (0.99 - 721.5)^2 +$$

$$(0.99 - 888.5)^2 + (0.99 - 430)^2 + (0.99 - 818.5)^2 \Big]$$

$$= \frac{1}{5} \Big[(-684.01)^2 + (-720.51)^2 + (-887.5)^2 + (-429.0$$

$$+ (-817.5)^2 \Big]$$

$$= 2627014 / 5$$

$$MSE = 525402$$

V Compute Cost function.

[Given $\lambda = 0.5$, $V - 3$ (No of features)]

$$C(V, \omega, I) = MSE - \lambda \cdot |V|$$

$$= 525402 - 0.5\,|3|$$

$$= 525402$$

Next Calculate the same for Subsets.

① $f_1, f_2 = c(V, \omega, I)$

② $f_1, f_3 = c(V, \omega, I)$

③ $f_2, f_3 = c(V, \omega, I)$

For subset f_1, f_2

$c(V, \omega, I)$

x_1	27	7.3
x_2	28	9.4
x_3	32	9.2
x_4	15	4.5
x_5	33	11.4

Repeat all the steps to calculate $c(V, \omega, I)$ by Considering only two features.

$$I(\omega) \quad x_1 = \frac{1}{1+e^{-(0.5(27)+0.2(7.3))}}$$

$$= \frac{1}{1+e^{-14.96}} = 0.999$$

$$x_2 = \frac{1}{1+e^{-(0.5(28)+0.2(7.4))}} = 0.9999$$

$$x_3 = \frac{1}{1+e^{-(0.5(32)+0.2(9.2))}} = 0.9999$$

$x4 = \dfrac{1}{1+e^{-(0.5(15)+0.2(4.5))}} = 0.9999$

$x5 = \dfrac{1}{1+e^{-(0.5(33)+0.2(11.4))}} = 0.9999$

Calculate f(wi) for this subset

$x1 = 10 + 12(27) + 15(7.3)$

$= 10 + 324 + 109.5 = 443.5$

$x2 = 10 + 12(28) + 15(7.4)$

$10 + 336 + 111 = 457$

$x3 = 10 + 12(32) + 15(9.2) = 10 + 384 + 138$

$= \cancel{538}\ 532$

$x4 = 10 + 12(15) + 15(4.5) = 10 + 180 + 67.5$

$= 257.5$

$x5 = 10 + 12(33) + 15(11.4) = 10 + 396 + 171$

$= 577$

Calculate MSE

$$MSE = \frac{1}{5}\left[(0.99-443.5)^2 + (0.99-457)^2 + (0.99-532)^2 + (0.99-257.5)^2 + (0.99-577)^2\right]$$

$= 216663.33$

$c(v,w,I) = 216663.33 - (0.5 * 2)$

$= 216662.34$

Similarly for second subset

			$I(w_i)$	$f(w_i)$
x_1	27	21	0.999	575.5
x_2	28	23	0.999	587.5
x_3	32	31	0.999	750.5
x_4	15	15	0.999	362.5
x_5	33	21	0.999	647.5

$$MSE = 273172.78$$

$$C(v, w, I) = 273172.78 - 0.5 * 2$$

$$= 273171.78$$

Inference :

By comparing the cost function of the subset, which subset is having the Maximum cost function values, those subset (features) are Important (i.e) Contributed more.