Lead Scoring Case Study

Q1.

Ans:

$$\ln\left(\frac{\hat{p}}{(1-\hat{p})}\right) = b_0 + b_1X_1 + b_2X_2 + \ldots + b_pX_p$$

As per above mathematical equation, logistic regression has linear equation. Where b1 and b2 are coefficient.

As per our model "model1" in Lead case study, we have some coefficient which shows weight of variable in model.

Variable	coef	z	[0.025	0.975]
Tags_Lost to EINS	8.8874	11.689	7.397	10.378
Tags_Closed by Horizzon	7.7602	10.002	6.24	9.281
Tags_Will revert after reading the email	3.8881	16.482	3.426	4.35
Tags_Busy	3.3903	10.18	2.738	4.043
Lead Source_Welingak Website	3.2282	3.938	1.622	4.835
Last Notable Activity_SMS Sent	2.6353	20.628	2.385	2.886
Lead Origin_Lead Add Form	2.0428	5.634	1.332	2.753
What is your current occupation_Working Professional	1.3694	4.277	0.742	1.997
Total Time Spent on Website	0.9228	18.317	0.824	1.022
Do Not Email	-1.4434	-6.309	-1.892	-0.995
Tags_Ringing	-1.6383	-4.745	-2.315	-0.962
const	-2.2592	-9.968	-2.703	-1.815
Tags_switched off	-2.3175	-3.824	-3.505	-1.13
Lead Quality_Not Sure	-3.0535	-22.853	-3.315	-2.792
Lead Quality_Worst	-3.6889	-4.141	-5.435	-1.943

As we can see coefficient of variables so on the bases of above evidence, we can say that below variables are top three which are contributing most towards the probability of a lead getting converted.

1. **Tags**
2. **Lead Source**
3. **Last Notable**

Q2:

Ans: Below Dummy variables are in the model which should be focused more to convert hot lead.

1. Tags_Lost to EINS
2. Tags_Closed by Horizzon
3. Tags_Will revert after reading the email

Q3:

Ans: In this case we need high "Sensitivity" and low "Specificity". We will take cutoff point as lower as satisfy business requirement. As business recommend that they want maximum lead as much as possible then we will take very low cutoff point and, in this case, False Positive become very high.

Model Strategy:

1. High Sensitivity
2. Low Specificity

Q4.

Ans:

As company said we don't need to call unless lead is very hot. So, in this case we will change our model prediction strategy. We will consider only those leads who all have very high probability to convert. We will consider **very high specificity** and **very low sensitivity** approach in model. So that we can achieve own purpose.

We will keep in mind that False positive should be very low so, that there is very minimum chance of wrong prediction such that business save their time to call cool lead.

Model Strategy:

1. Low Sensitivity
2. High Specificity