

Subjective Questions – Lead Score Assignment
Krishnan Vybhava Srinivasan and Nithya V

1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?

Response: The higher the co-efficient the higher impact it has on the resultant variable and hence, the probability of it being converted (other things remaining constant). Based on the above, the top three variables, with highest positive correlation are:

Variable	Coefficient
Tags - Lost to EINS	9.2288
Tags_Closed by Horizzon	8.2464
LeadSource_Welingak Website	4.3047

2. What are the top 3 categorical/dummy variables in the model which should be focused the most on in order to increase the probability of lead conversion?

The Dummy variables that has highest positive coefficient are as below:

Dummy Variable	Original Categorical Variable	Coefficient
Tags - Lost to EINS	Tags	9.2288
Tags_Closed by Horizzon	Tags	8.2464
LeadSource_Welingak Website	LeadSource	4.3047
Tags_Will revert after reading the email	Tags	3.8485
Tags_Busy	Tags	3,6020
Lastnoteableactivity_SMS Sent	Lastnoteableactivity	2.2747
occupation_Working Professional	Occupation	1.3313

Based on the above the Top Three Categorical Values, with relevant Levels are:

Original Categorical Variable	Dummy Variable
Tags	Lost to EINS, Closed by Horizzon, Will revert after reading the email, Busy
LeadSource	Welingak Website
Lastnoteableactivity	SMS Sent
Occupation	Working Professional

Subjective Questions – Lead Score Assignment

Krishnan Vybhava Srinivasan and Nithya V

3. X Education has a period of 2 months every year during which they hire some interns. The sales team, in particular, has around 10 interns allotted to them. So during this phase, they wish to make the lead conversion more aggressive. So they want almost all of the potential leads (i.e. the customers who have been predicted as 1 by the model) to be converted and hence, want to make phone calls to as much of such people as possible. Suggest a good strategy they should employ at this stage.

Response: In this case the Company wants to make sure we predict all the leads which are likely to be converted. In other words, the strategy would be to increase the sensitivity metrics (also called the **true positive rate**, the **recall**, or **probability of detection of conversion**)

	prob	accuracy	sensi	speci	preci
0.0	0.0	0.385136	1.000000	0.000000	0.385136
0.1	0.1	0.762715	0.974652	0.629962	0.622617
0.2	0.2	0.844906	0.937858	0.786684	0.733610
0.3	0.3	0.919855	0.857727	0.958771	0.928730
0.4	0.4	0.919698	0.856909	0.959027	0.929078
0.5	0.5	0.920170	0.855683	0.960563	0.931464
0.6	0.6	0.920170	0.850777	0.963636	0.936122
0.7	0.7	0.917651	0.840965	0.965685	0.938841
0.8	0.8	0.872304	0.687244	0.988220	0.973364
0.9	0.9	0.856243	0.639411	0.992061	0.980564

Currently our model, has a sensitivity/recall of 85.8% at probability cut off of 0.30. Given for 2 months every year we have bandwidth to make more calls and objective is to convert all people predicted by model to be converted and has ability to make as many phone calls as possible. We suggest that the sensitivity/recall % should be increased even if there is a drop-in accuracy and specificity score.

Hence, we propose a reduction in probability cut off from 0.30 to 0.10. In other words, classify any leads with conversion probability in excess of 0.10 as Hot Leads. As at this range the recall % will be 97% and accuracy at 76%. Although the precision % will drop from 93% to 63%

Subjective Questions – Lead Score Assignment

Krishnan Vybhava Srinivasan and Nithya V

4. Similarly, at times, the company reaches its target for a quarter before the deadline. During this time, the company wants the sales team to focus on some new work as well. So during this time, the company's aim is to not make phone calls unless it's extremely necessary, i.e. they want to minimize the rate of useless phone calls. Suggest a strategy they should employ at this stage.

Response: During the period when the Company reaches its quarterly deadline, it would like the sales team to focus on new work and hence would like to work on leads where it is extremely necessary.

Hence, the expectation at this stage is to ensure Precision Rate is high, which means Probability that a predicted 'Yes' is actually a 'Yes'

Currently our model, has a precision is 93 % at a probability cut off of 0.30. Given sales team would like to work on leads where it is extremely necessary.. We suggest that the precision % should be increased during these stages to ensure they work on calls which will for sure result in conversion.

	prob	accuracy	sensi	speci	preci
0.0	0.0	0.385136	1.000000	0.000000	0.385136
0.1	0.1	0.762715	0.974652	0.629962	0.622617
0.2	0.2	0.844906	0.937858	0.786684	0.733610
0.3	0.3	0.919855	0.857727	0.958771	0.928730
0.4	0.4	0.919698	0.856909	0.959027	0.929078
0.5	0.5	0.920170	0.855683	0.960563	0.931464
0.6	0.6	0.920170	0.850777	0.963636	0.936122
0.7	0.7	0.917651	0.840965	0.965685	0.938841
0.8	0.8	0.872304	0.687244	0.988220	0.973364
0.9	0.9	0.856243	0.639411	0.992061	0.980564

Hence, we propose a increase in probability cut off from 0.30 to 0.90. In other words, classify any leads with conversion probability in excess of 0.90 as Hot Leads. As at this range the precision % will be 98% and accuracy at 86%. Although the Recall % will drop from 92% to 64%