

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

Answer:

Following are three top variables which are contributing towards probability of a lead being converted to customer:

- **Total Time Spent on Website:**
 - Leads who are spending more time on the website are having higher chance of getting converted
- **Lead_Origin_Lead Add Form:**
 - If lead origin is lead add form there is more likely to get converted.
- **Occupation_Working Professional:**
 - If the lead customer is working professional, then there is more chance that the customer will be converted.

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?

Answer:

Following are three categorical dummy variables which need to be focussed:

- **Lead_Origin_Lead Add Form:**
 - If lead origin is lead add form there is more likely to get converted.
- **Last_Ntbl_Activity_Olark Chat Conversation:**
 - As per the model last notable activity done by the student / customer on olark chat conversion which negatively impacting the lead conversion.
- **Occupation_Working Professional:**
 - If the lead customer is working professional, then there is more chance that the customer will be converted.

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.

Answer:

X Education has 10 interns allotted to them and hence during this phase they need to be calling all potential leads and try to make maximum conversions. So customer which are predicted as 1 by the model will surely get converted. So we need to make sure to design the model which has high true positive rate (i.e. higher sensitivity).

Eg. X Education can consider cutoff probability of 0.20, i.e. contact all customer with lead score > 20. In this case sensitivity metrics will be as high as 91%, with decent accuracy of 75%.

Probability	Accuracy	Sensitivity	Specificity	Precision	Recall
0.00	0.385136	1.000000	0.000000	0.385136	1.000000
0.05	0.490474	0.993868	0.175160	0.430113	0.993868
0.10	0.626201	0.975470	0.407426	0.507660	0.975470
0.15	0.712959	0.946034	0.566965	0.577778	0.946034
0.20	0.754369	0.918234	0.651729	0.622851	0.918234
0.25	0.798142	0.884710	0.743918	0.683944	0.884710
0.30	0.808534	0.864677	0.773367	0.705000	0.864677
0.35	0.812785	0.821341	0.807426	0.727635	0.821341
0.40	0.811841	0.772281	0.836620	0.747527	0.772281
0.45	0.811211	0.732216	0.860691	0.767024	0.732216
0.50	0.811998	0.695830	0.884763	0.790892	0.695830
0.55	0.808219	0.659035	0.901665	0.807615	0.659035
0.60	0.806960	0.630008	0.917798	0.827605	0.630008
0.65	0.800661	0.594440	0.929834	0.841435	0.594440
0.70	0.780507	0.517171	0.945455	0.855886	0.517171
0.75	0.772319	0.479150	0.955954	0.872024	0.479150
0.80	0.764447	0.440311	0.967478	0.894518	0.440311
0.85	0.742875	0.365495	0.979257	0.916923	0.365495
0.90	0.718942	0.293949	0.985147	0.925354	0.293949
0.95	0.680680	0.182339	0.992830	0.940928	0.182339

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.

Answer:

Since company wants to be minimize and avoid useless calls, so model should predict minimum false positives..

So X Education should consider cutoff probability of 0.75 or 0.80. So contact all the customers whose lead scores are more then 75. In this case precision should be higher like 87.2%. With decent accuracy as 77%. So its fine to have lower sensitivity; and in this case it is 47.9%. We may not convert to have true positives but it will serve agenda of company to make calls when it is only needed.

Probability	Accuracy	Sensitivity	Specificity	Precision	Recall
0.00	0.385136	1.000000	0.000000	0.385136	1.000000
0.05	0.490474	0.993868	0.175160	0.430113	0.993868
0.10	0.626201	0.975470	0.407426	0.507660	0.975470
0.15	0.712959	0.946034	0.566965	0.577778	0.946034
0.20	0.754369	0.918234	0.651729	0.622851	0.918234
0.25	0.798142	0.884710	0.743918	0.683944	0.884710
0.30	0.808534	0.864677	0.773367	0.705000	0.864677
0.35	0.812785	0.821341	0.807426	0.727635	0.821341
0.40	0.811841	0.772281	0.836620	0.747527	0.772281
0.45	0.811211	0.732216	0.860691	0.767024	0.732216
0.50	0.811998	0.695830	0.884763	0.790892	0.695830
0.55	0.808219	0.659035	0.901665	0.807615	0.659035
0.60	0.806960	0.630008	0.917798	0.827605	0.630008
0.65	0.800661	0.594440	0.929834	0.841435	0.594440
0.70	0.780507	0.517171	0.945455	0.855886	0.517171
0.75	0.772319	0.479150	0.955954	0.872024	0.479150
0.80	0.764447	0.440311	0.967478	0.894518	0.440311
0.85	0.742875	0.365495	0.979257	0.916923	0.365495
0.90	0.718942	0.293949	0.985147	0.925354	0.293949
0.95	0.680680	0.182339	0.992830	0.940928	0.182339