

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

Sol. Based on the coefficient values from below screenshot, the following are the top three variables that contribute most towards the probability of a lead getting converted.

1. Lead Source
2. What is your Current occupation.
3. Total Time Spent.

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?

Sol. Based on the coefficient values from below screenshot, the following are the top three variables that contribute most towards the probability of a lead getting converted.

1. Lead Source_Welingak Website.
2. Lead Source_Reference.
3. What is your current occupation_Working Professional

| | coef |
|--|---------|
| const | -0.5146 |
| Total Time Spent on Website | 1.0668 |
| A free copy of Mastering The Interview | -0.4331 |
| Specialization_Hospitality Management | -1.0055 |
| Specialization_Other | -0.4918 |
| Lead Source_Olark Chat | 1.0712 |
| Lead Source_Reference | 3.4227 |
| Lead Source_Welingak Website | 5.5705 |
| What is your current occupation_Unknown | -1.2425 |
| What is your current occupation_Working Professional | 2.4345 |

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.

Sol:

- High sensitivity implies that our model will correctly identify almost all leads who are likely to Convert. It will do that by over-estimating the Conversion likelihood, i.e. it will misclassify some non-Conversion cases as Conversions.
- Now, since X Education has more man-power for these 2 months and they wish to make the lead conversion more aggressive by wanting almost all of the potential leads, we can choose a lower threshold value for Conversion Probability.
- To make the sales aggressive, the company may contact all the leads which have a conversion probability Lead Score >0.2 .
- When we choose a lower Cutoff value to increase the Sensitivity of our model which will then Predict the cases with Lower probabilities as Converted.

| | final_predicted | Lead_Score |
|------|-----------------|------------|
| 4269 | 0 | 22.79 |
| 2376 | 1 | 81.33 |
| 7766 | 1 | 74.85 |
| 9199 | 1 | 29.32 |
| 4359 | 1 | 87.69 |
| 9186 | 1 | 43.42 |
| 1631 | 1 | 62.67 |
| 8963 | 0 | 19.34 |
| 8007 | 0 | 11.61 |
| 5324 | 0 | 24.11 |
| 2558 | 1 | 29.68 |
| 8973 | 0 | 22.60 |
| 2935 | 1 | 67.88 |

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.

Sol. High Specificity implies that our model will correctly identify almost all leads who are not likely to Convert. It will do that at the cost of losing out some low Conversion rate risky leads to the competition, i.e. it will misclassify some Conversion cases as non-Conversions.

Therefore, since X Education has already reached its target for a quarter and doesn't want to make phone calls unless it's extremely necessary, i.e. they want to minimize the rate of useless phone calls, we can choose a higher threshold value for Conversion Probability.