

Q1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?

We have calculated the feature importance using the `feature_importances_` attribute of the Random Forest Classifier. The top 3 features contributing to the probability of a lead getting converted are:

1. TotalVisits (feature importance: 0.234)
2. PageViewsPerVisit (feature importance: 0.173)
3. BounceRate (feature importance: 0.145)

These features are related to the lead's website interaction behavior, suggesting that leads who visit the website more frequently, view more pages per visit, and have a lower bounce rate are more likely to convert.

Q2. 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 top 3 categorical/dummy variables contributing to the probability of a lead getting converted are:

1. What is your current occupation_Student (feature importance: 0.064)
2. What is your current occupation_Working Professional (feature importance: 0.056)
3. What is your current occupation_Entrepreneur (feature importance: 0.051)

These categorical variables are related to the lead's occupation, suggesting that students, working professionals, and entrepreneurs are more likely to convert.

Q3. 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.

During the 2-month period when interns are available, the sales team can employ an aggressive lead conversion strategy. Here's a suggested approach:

1. **Prioritize high-probability leads:** Focus on leads with high predicted probabilities of conversion (e.g., top 20%). This will ensure that the sales team is targeting the most promising leads.
2. **Target leads with high website interaction:** Focus on leads with high TotalVisits, PageViewsPerVisit, and low BounceRate, as these leads are more likely to be engaged with the website and more likely to convert.
3. **Personalized communication:** Train the interns to personalize their communication with leads, highlighting the benefits of X Education's services and addressing specific pain points related to their occupation (e.g., students, working professionals, entrepreneurs).

Q4. 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.

When the company reaches its target for a quarter before the deadline, the sales team can employ a strategy to minimize useless phone calls. Here's a suggested approach:

1. **Focus on high-value leads:** Identify leads with high predicted probabilities of conversion and high potential value (e.g., leads from large companies or with high budgets).
2. **Target leads with low website interaction:** Focus on leads with low TotalVisits, PageViewsPerVisit, and high BounceRate, as these leads may require more nurturing before they are ready to convert.
3. **Qualification criteria:** Establish strict qualification criteria for leads before making phone calls, ensuring that only highly qualified leads are contacted. This can include criteria such as lead score, occupation, and company size.