

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

The three variables that contribute the most to the positive conversion of a lead in order of priority are:

- i. Tags\_Closed by Horizzon with a Positive coeff of 9.2007
- ii. Lead\_Profile\_Others with a Positive coeff of 3.9186
- iii. Tags\_Will revert after reading the email with a Positive coeff of 4.1022

*Note: Variable 2, Lead Profile Others represents a variable created based on data cleaning. The residual categories under this feature were aggregated into the Others category as they form less than 2% of the entire data set. (Review the PPT- Data cleaning section for further details)*

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?

All three variables above defined in question 1 are categorical variables. The highest degree of conversion could be attributable to the three variables above.

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.

Based on the limited information deduced from the model, the following strategies could be laid down/formulated:

- a. If the Tag representing closed by Horizzon represents an agent, we should review commercial arrangements with them to confirm a higher conversion and customer base that they should cater to.
- b. Lead profile represents the residual category of leads which is an aggregation of all categories of leads less than 2% of the total count. The smaller categories of other lead profiles have shown a higher conversion. When the data is reviewed closely, these relate to categories other than that identified as potential lead (which is essentially the mode value in the data set). The strategy could be that the company investigates the sources that help the data be tagged as a potential lead. If this variable was statistically significant or important to the data set, it would have been included in the RFE test. Hence, this variable or identification being categorized as a 'potential lead' may not itself be a potential lead as the other categories have a higher relation with respect to conversion.
- c. Where customers have committed to reading the email, the company should focus on a quick turnaround to effectively communicate with the customer. Support can be offered to the customer to close gaps in understanding, or just prevent the lead from dying out due to loss of interest, or the customer getting busy, or any other unforeseen impediments to the lead conversion.

- d. Leads generated through add forms are effective. Once the company has realized that the customer has approached the lead add form, a quick turnaround through calling can be adopted to validate that the customer is interested. The assumption is that once the lead form is generated through advertisements or any other sources, the customer has already expressed some interest. This interest can be effectively converted through personal and positive customer interaction along with the interest within the product offering being retained/strengthened.
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.
- Based on the limited information deduced from the model, the following strategies could be laid down/formulated:
- a. Once customers are sent an SMS, follow-up communication can be triggered through a detailed email to convert the customer. As observed above, both the SMSs and emails that are read have led to an effective conversion strategy.
  - b. As mentioned above, Horizzon, which either represents a channel of generated leads or a certain service provider, has a high success rate with respect to lead conversion. Commercial arrangements can be reviewed as well as management discussions were done to ensure that this channel of lead generation/ creation has traction or growth.
  - c. Where the customer is tagged as ringing or switched off, this could represent a busy customer. An attempt can be made to review the incremental measures placed for such customers as the lead conversion may not be effective. It may end up being a waste of resources that could have been probably differently utilized.
  - d. Customer activity through time spent on websites or where the customer has some recorded 'Last activity' through a review of data available. Some incremental resources expended to increase online interaction or send notifications or system pings based on the last activity recorded may result in better conversions.