

## **Assignment Subjective Questions**

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

Top 3 variables which contribute most towards the probability of a lead getting converted are: -

- Tags Conversion Chances [Category HIGH CONVERSION CHANCES]
  - This is derived column based on 'Tags' in original dataset, and category 'LOW\_CONVERSION\_CHANCES' in column represents tag values – (1) Will revert after reading the email, (2) Interested in Next batch, (3) Lost to EINS, (4) Closed by Horizzon, (5) Lateral student.
  - This variable has positive effect towards the probability of a lead getting converted.
- Last Notable Activity [Category 'SMS Sent']
  - This variable has positive effect towards the probability of a lead getting converted.
- Lead Source [Category 'Direct Traffic']
  - This variable has negative effect towards the probability of a lead getting 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?**

To increase the probability of lead conversion, only variables with positive coefficient values are considered. Top 3 categorical/dummy variables in the model to be focused the most on in order to increase the probability of lead conversion: -

- Tags Conversion Chances [Category HIGH CONVERSION CHANCES]
  - Historical responses by leads like 'Will revert after reading the email', 'Interested in Next batch', and indications by sales team based on interaction with leads like 'Lost to EINS', 'Closed by Horizzon', 'Lateral student', provide very good conversion rate at later stage.
- Last Notable Activity [category 'SMS Sent']

- SMSs have great impact on leads.
- Total Time Spent on Website:
  - More time spent by lead, indicates higher chances of lead to be converted to customer.

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.

X Education in to prioritize the factors that have positive effect on probability of lead conversion, and should avoid the factors that have negative effect. A good strategy to get the maximum outcome would be as follows: -

- Model Adjustment and Enhancement:
  - Decrease the model cut-off probability to have higher 'Recall', which will cover most of leads that have any chance to be converted. Although the result will include leads where chances of conversion is low, but it's mentioned that good manpower is available, so its not adequate to leave any lead that may be converted to customer.
  - Rebuild model frequently to use latest statistic and better results.
  - Making model from scratch, that may include exploring algorithms other than logistic regression, and use best model for better results.
- Additional Strategies to increase the reachability:
  - Target the leads who are tagged 'Will revert after reading the email', 'Interested in Next batch', 'Lost to EINS', 'Closed by Horizzon', 'Lateral student', have good conversion. So, X Education should prioritize such leads.
  - The audience/leads that are approached through SMS have provided positive outcomes, so it increasing reach to audience/leads through SMS should be increased.

- It is observed that leads that spend good time on X Education's website have higher chances of converting to customer. So, X Education can work on enhancing feature and information on their website, and make it more interactive.
- Leads that open X Education Emails are also good candidates to target, so title for these Emails can be made more innovative, make user to open to know more about programs.

**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.**

To minimize the rate of useless phone calls, company may use following strategy:

- **Model Adjustment and Enhancement:**

- Increase the model cut-off probability to have higher 'Precision', which will return more relevant results, rather than irrelevant results.
- Rebuild model frequently to use latest statistic and better results.
- Making model from scratch, that may include exploring algorithms other than logistic regression, and use best model for better results.

- **Additional Strategies to minimize useless calls:**

- Specifically target the leads that are already tagged to have high chances of conversion.
- Follow-up through automated SMS service, and also try to reach new focused audience.
- Enrich Website to support chatbot and multi-level questioning to get more in-depth information of landed audience and identify only highly interested audience.
- Whenever phone service is used than choose the lead with higher 'Lead Score' at any stage.