1. Which are the top three variables in your model which contribute most 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?
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.
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.

# Answers to Subjective Questions

### 1. **Top Three Variables Contributing to Lead Conversion**

Based on the logistic regression model and feature importance analysis, the top three variables contributing most to lead conversion are:

* **Total Time Spent on Website**: Leads spending more time on the website are significantly more likely to convert.
* **Page Views Per Visit**: Higher engagement with multiple pages correlates with higher conversion probability.
* **Lead Source - Google**: Leads sourced from Google show a higher likelihood of conversion.

### 2. **Top Three Categorical/Dummy Variables to Focus On**

The top three categorical variables identified as crucial for lead conversion are:

* **Lead Source - Google**: Indicates that Google is a critical source of high-quality leads.
* **Last Activity - Email Opened**: Demonstrates that email engagement is a strong predictor of conversion.
* **Specialization - Finance Management**: Suggests that targeting professionals in finance management could yield better results.

### 3. **Strategy for Aggressive Lead Conversion During Intern Period**

To maximize lead conversion during the intern period, X Education should:

* Focus on leads with a predicted probability of conversion greater than **0.8**.
* Assign interns to follow up with "warm" leads (predicted probabilities between 0.6 and 0.8) through personalized communication.
* Implement automated email campaigns to engage "cold" leads and potentially warm them up for future conversion efforts.

### 4. **Strategy to Minimize Useless Calls During Target Met Period**

To reduce unnecessary calls when quarterly targets are met, X Education should:

* Set a higher threshold for lead follow-ups (e.g., predicted probability > **0.9**).
* Focus on leads with recent high-engagement activities, such as "Form Submitted" or "Video Viewed."
* Leverage automated email campaigns for low-probability leads to maintain engagement without direct intervention.

### Additional Recommendations

* **Data Collection Improvements**: Address missing data in key features like Specialization and Tags by refining data collection methods.
* **Model Refinement**: Consider exploring ensemble models like Random Forest or Gradient Boosting for improved accuracy.
* **Operational Enhancements**: Develop a scoring dashboard to enable real-time lead prioritization and train the sales team to utilize model insights effectively.