Assignment Subjective Questions’ Answers

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

**Ans:** To determine this, we look at the **coefficients (odds ratio) in** *res.summary()* from the logistic regression model. The variables with the highest absolute coefficients (positive or negative) are the most important.

From the model, the top 3 variables likely are:

* + 1. **Total Time Spent on Website** – Higher time spent usually means higher engagement.
    2. **Last Activity (e.g., Email Opened, Page Visited)** – Certain activities indicate strong interest.
    3. **Lead Source (e.g., Organic Search, Google Ads, Referral Sites)** – Some sources convert better than others.

# 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?

**Ans:** From *list(zip(X\_train.columns, rfe.support\_, rfe.ranking\_)),* we can identify categorical variables. Likely candidates include:

* + 1. **Lead Source (Google, Direct Traffic, etc.)** – Some perform better. Prioritize high-converting sources.
    2. **Last Notable Activity (e.g., Email Opened, SMS Sent)** – Target users with high engagement.
    3. **Specialization (e.g., Finance, Marketing, IT)** – Certain specializations may be more likely to convert.

# 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.

**Ans:** During the intern hiring phase, X Education should lower the model's probability threshold (e.g., from 0.5 to 0.3) so that more leads are classified as potential conversions. This means they’ll reach out to more prospects, increasing their chances of engagement.

To maximize success, they should prioritize leads who show strong interest, such as those with high Total Time Spent on Website and past engagement (e.g., opened emails, interacted with chat). These leads are more likely to convert if contacted quickly.

Interns should be assigned to immediate follow-ups, ensuring that high-scoring leads are contacted while their interest is still fresh. This fast, data-driven approach will help maximize conversions during this period.

# 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.

**Ans:** When X Education has already met its quarterly target, the goal is to **reduce unnecessary calls** and focus only on the leads with the highest chances of conversion.

To do this, they should **increase the probability threshold** (e.g., from **0.5** to **0.7**), ensuring that only leads with a strong likelihood of converting are classified as high-priority. This helps the sales team focus on the **most valuable leads** instead of wasting time on low-potential ones.

Priority should be given to leads with **high Total Time Spent on Website** and positive engagement history (e.g., recent email interactions, chat inquiries, or call requests).

For **borderline leads** (who show some interest but don’t meet the higher threshold), instead of calling, the company can **use automated emails or SMS sequences** to keep them engaged without spending too many resources. This way, the sales team can stay productive while minimizing effort on less promising prospects.