**Lead Score Analysis**

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

Ans.

**Top 3 Variables Influencing Lead Conversion**

1. **Lead Origin:** Leads generated through the **Lead Add Form** have the highest conversion rate. This indicates that prospects who engage via Lead Add Forms are more likely to convert.
2. **Lead Source:** The **Welingak Website** and **Referrals** yield the highest conversion rates. Leads coming from these sources tend to be of higher quality, suggesting that focusing on these channels can enhance lead conversion.
3. **Last Activity:** Leads who engage through **Olark Chat Conversations** show the highest conversion potential. Active interactions via chat and telephonic conversations significantly impact conversions, emphasizing the importance of timely follow-ups and improved chatbot engagement.

**2. What are the top 3 categorical/dummy variables in the model which should be focused the most on to increase the probability of lead conversion?**

Ans.

TOP 3 Variables

1. **Lead Origin**, wherein Lead Add Form has highest conversion

Leads who have responded/ or engaged through Lead Add Forms have had a higher chance of getting converted

1. Lead Source, wherein Welingak Website and reference has highest conversion.

The conversion rate is high for the leads that the X education get from Welingak Website and references, we can work on these two sources to get good quality leads

1. Last Activity, wherein Olark Chat Conversation has highest conversion

The leads which are actively interacting with the X\_education through Olark chat and telephonic conversation have good conversion. This suggests follow-ups from chat interactions are important → Improve chatbot engagement.

**3. X Education has a period of 2 months every year during which they hire some interns. The sales team 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 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.

**Objective:**  
During the 2-month internship period, the company wants to maximize lead conversions, meaning they are willing to call as many high-potential leads as possible.

**Strategy:**

* Lower the classification threshold (e.g., from 0.5 to 0.3 or lower) so that more leads are classified as 1 (potential converters).
* This increases recall, meaning fewer potential customers are missed.
* Since interns are available, the cost of extra phone calls is low, so even if some low-probability leads are included, it won’t be a major issue.

**Implementation:**

1. Adjust the threshold dynamically:

threshold = 0.3 # Reduce threshold for aggressive lead conversion

y\_pred\_adjusted = (y\_predict\_proba >= threshold).astype(int)

1. **Prioritize high-importance features for segmentation.**
   * Focus on leads coming from Lead Add Form, References & Welingak Website, as they have the highest conversion potential.
   * Use Last Activity (Olark Chat, Phone Conversation, etc.) to prioritize.
   * If available, rank leads by probability and start calling from the highest scores down.

**Expected Outcome:**

Higher lead conversions but also a higher number of total calls made (some wasted effort is acceptable here).

**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.**

**Ans.**

**Objective:**  
Once the company has met its target, the goal shifts to avoiding unnecessary calls and minimizing false positives (people predicted as leads but don’t actually convert).

**Strategy:**

* Increase the classification threshold (e.g., from 0.5 to 0.7 or higher).
* This improves precision, ensuring that only the most confident leads receive calls.
* Leads with low probability (even if predicted as 1 at a lower threshold) are ignored.

**Implementation:**

1. Adjust the threshold dynamically:

threshold = 0.7 # Increase threshold to avoid false positives

y\_pred\_adjusted = (y\_predict\_proba >= threshold).astype(int)

1. **Only call leads from the highest-scoring categories**
   * Leads from Welingak Website & Lead Add Form → Higher priority
   * Avoid calling leads from low-importance sources
   * Check for past interaction history (e.g., people who engaged in Olark Chat may still be worth calling)

**Expected Outcome:**

* Fewer phone calls made, but each call has a higher chance of conversion.
* Saves effort & improves efficiency for the sales team.