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

Top Three Features Impacting Lead Conversion

1. Tags (Coefficient: 3.30, Positive Impact)
    - This is the most significant feature influencing conversion.
    - Likely represents categorized labels for leads, such as "Hot Lead" or "Interested Prospect."
    - Leads with certain tags have a higher probability of conversion.
2. Lead Origin - Lead Add Form (Coefficient: 2.81, Positive Impact)
    - Indicates that leads coming from the Lead Add Form (probably a website inquiry form) have a strong positive impact on conversion.
    - This suggests that directly captured leads are more likely to convert.
3. Lead Profile - Student of Some School (Coefficient: -2.04, Negative Impact)
    - A negative coefficient means leads labelled as "Student of Some School" are less likely to convert.
    - Possibly, students may be exploring options but are not serious buyers.

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

Top categorical variables that have the strongest positive influence on lead conversion (highest positive coefficients) are:

Rank	Feature Name	Coefficient	Impact
1	Tags	3.302	Strongest positive impact
2	Lead Origin - Lead Add Form	2.811	High impact, focus on optimizing this lead source
3	Lead Source - Welingak Website	1.885	Important lead source, should be prioritized

Insights

1. Tags (3.30) → Strongest Predictor of Conversion
    - Leads with specific tags are highly likely to convert.
    - Action: Identify which tags perform best and target similar leads.
2. Lead Origin - Lead Add Form (2.81) → Highly Effective Lead Source
    - Leads coming from the Lead Add Form convert more often.
    - Action: Optimize and promote this form more aggressively.
3. Lead Source - Welingak Website (1.88) → Strong Lead Source
    - Leads generated from the Welingak Website are valuable.
    - Action: Invest in marketing strategies that increase traffic to this website.

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.**

   a) Prioritize High-Probability Leads for Calling
   - Sort leads by predicted probability (from predict_proba output).
   - Segment leads:
     - High confidence (0.80 - 1.00 probability) → Call immediately
     - Moderate confidence (0.60 - 0.79 probability) → Call in the second wave
     - Low confidence (below 0.60 probability) → Use alternative channels (emails, SMS)

   b) Focus on Key Influencing Factors for Calls

   From the model, phone conversations significantly improve conversion chances (Last Notable Activity - Had a Phone Conversation has a high coefficient).
   - Ensure that all high-priority leads receive at least one phone call.
   - Interns should follow a structured script:
     - Highlight key selling points of the course.
     - Address common objections.
     - Offer time-limited incentives (e.g., discounts, free trials).

   c) Optimize Calling Strategy with Interns
   - Allocate interns based on lead priority:
     - 5 interns → Call high-confidence leads first
     - 5 interns → Call moderate-confidence leads next
   - Use CRM to track progress:
     - Mark leads who did not pick up for a callback attempt.
     - Send follow-up emails/SMS if they do not respond.

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 is extremely necessary, i.e. they want to minimize the rate of useless phone calls. Suggest a strategy they should employ at this stage.**

   Focus Only on High-Quality Leads

Instead of calling every predicted lead, filter out those who are least likely to convert based on the model's probability scores.

Steps to Implement:

- Set a probability threshold (e.g., 0.90+ confidence only).
- Eliminate leads with low engagement (e.g., no previous interactions).
- Prioritize past interactions like email opens, website visits, and responses.

Implementation: