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

Ans: From the regression coefficients, the top 3 variables with the **largest positive coefficients** contribute the most towards the probability of lead conversion:

1. **Lead Source\_Welingak Website** (coef = 6.0484): Strongest predictor for lead conversion.
2. **Lead Source\_Reference** (coef = 3.8284): Indicates referrals are highly impactful.
3. **Last Activity\_Had a Phone Conversation** (coef = 2.8997): Suggests that personal phone conversations significantly boost conversions.

As the features were selected via RFE, the only numerical variable in the model contributing significantly is **Total Time Spent on Website** (coef= 1.1139) which denotes student engagement is crucial for online education system.

1. 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: The top 3 categorical/ dummy variables that are actionable and can guide strategies to increase lead conversion:

**Lead Source\_Welingak Website**:

* Focus on optimizing the content and user experience of the Welingak website to attract more leads.

**Lead Source\_Reference**:

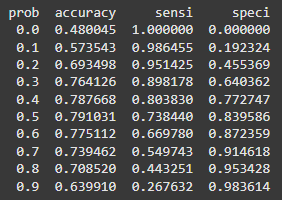
* Leverage referral programs to encourage existing leads/customers to refer others.

**Last Activity\_Had a Phone Conversation**:

1. Emphasize phone conversations as part of the follow-up strategy for potential leads.
2. 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 2-month intern phase, the company wants to convert as many leads as possible (high recall/sensitivity):

* **Optimize for Recall (Sensitivity)**: Adjust the decision threshold of the model to a lower value. This increases the proportion of predicted positives, ensuring that almost all potential leads are classified as 1 (positive) as lowering the threshold increases the number of predicted conversions, reducing the chance of missing potential leads.



Suggested Strategy:

1. Use the predicted probabilities from the model to sort leads by likelihood of conversion.
2. Assign high-priority leads (with the highest predicted probabilities) to experienced salespeople.
3. Use the interns for leads with moderate predicted probabilities to maximize outreach.
4. Focus on leads associated with the strongest predictors (e.g., Welingak Website, Reference, Phone Conversations).

1. 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 the company has already achieved its target, it wants to minimize unnecessary phone calls (low precision):

Optimize for Precision: Adjust the decision threshold to a higher value. This ensures that only leads with a high likelihood of conversion are classified as 1 (positive), reducing unnecessary calls to unlikely leads.

Suggested Strategy:

1. Increase the decision threshold to prioritize high-probability leads only.
2. Focus only on leads with strong positive predictors (Lead Source\_Welingak Website, Lead Source\_Reference, Had a Phone Conversation).
3. Use targeted follow-ups for leads with activities that are most likely to lead to conversion (e.g., phone conversations or SMS sent).