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

As part of model building, we used 'Recursive Feature Elimination (RFE)' to find the top variables, contributing most to the probability of a lead getting converted.
According to the model, '**Total Time Spent on Website', 'Lead Add Form' and 'Last Activity_SMS Sent'** were the top 3 variables.

- **Total Time Spent on Website**: We observed that leads who spent more time browsing the website were more likely to convert to learners. The website could be made more engaging and interactive to maintain and improve lead conversion rate.
- **Lead Add Form:** Leads who filled out the form and provided their details were more likely to convert to learners. Sales team could focus more on these leads for higher conversion. Also, the form could be kept as short as possible with easy access through the website and other channels.
- **Last Activity_SMS Sent:** Lead conversion via SMS marketing reflects that the strategy to keep leads engaged, as part of follow up, has been successful. This strategy could be maintained and improved by providing interested leads an easily accessible point of contact, for further course information via SMS.

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?

Based on the created model, the categorical variables that impact lead conversion the most are

- **Lead Source_Direct Traffic:** Leads who directly visit X Educations website are more likely to convert as they already have an interest in furthering their education goals.
- **Lead Source_Welingak Website**: Leads generated through Welingak website have a high conversion rate.
- **Tags_Closed by Horizzon:** Leads closed by Horizzon have a low conversion rate. Understanding requirements of leads originating from this source could yield better results.

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.

During the intern hiring period, the sales team should prioritize reaching out to all predicted hot leads (Lead Score > 85).

[297]:
```
hot_leads=y_train_pred_final.loc[y_train_pred_final["Lead_Score"]>=85]
hot_leads
```

[297]:

	Converted	Converted_prob	Predicted	Final_Predicted	Lead_Score
1667	1	0.988475	1	1	99
2714	1	0.997064	1	1	100
4881	1	0.924203	1	1	92
1067	1	0.959763	1	1	96
2197	1	0.959763	1	1	96
...	...	...	...	...	...
1994	1	0.959763	1	1	96
2543	1	0.987208	1	1	99
7993	1	0.998497	1	1	100
84	1	0.940215	1	1	94
5825	1	0.998330	1	1	100

1825 rows × 5 columns

A possible strategy is:

- Increase phone call follow-ups for these leads.
- Offer limited-time discounts or promotions.
- Use a multi-touch approach (emails + calls + retargeting ads) to increase engagement.

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.

When the company has already reached its target, the focus should shift to minimizing unproductive calls. A recommended strategy is:

- Set a higher Lead Score threshold (e.g., 90+ instead of 85+).
- Prioritize only those leads that have engaged recently (e.g., visited the website in the last 7 days).
- Focus on automated email follow-ups rather than phone calls to maintain engagement without using excessive resources.