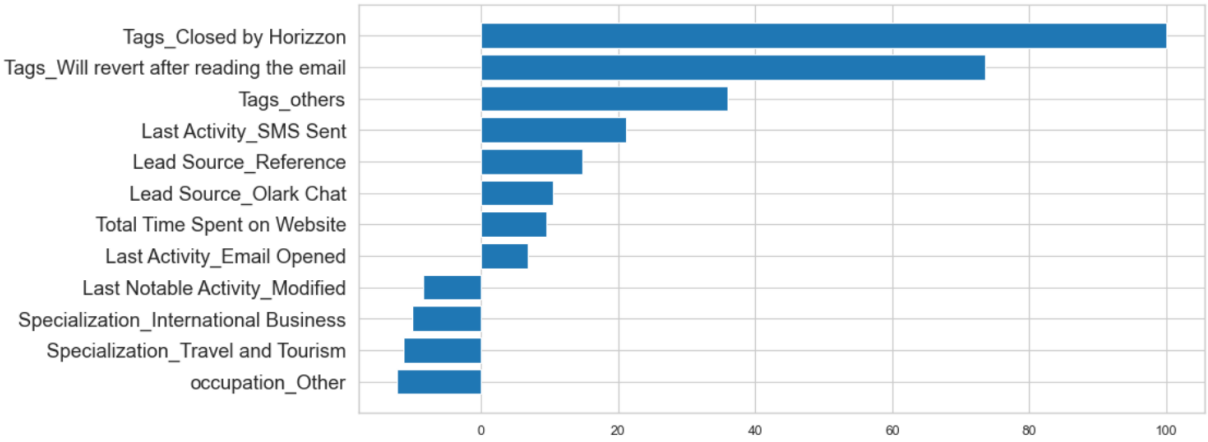
1. **Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?**

Ans- Twelve features were selected as the most significant in predicting the conversion:

Features having positive impact on conversion probability in decreasing order of impact basis model are:
1. Tags_Closed by Horizzon
2. Tags_Will revert after reading the email
3. Tags_others

But features having positive impact on conversion probability in decreasing order of impact basis business intuition are:
1. Tags_Closed by Horizzon
2. Tags_Will revert after reading the email
3. Last Activity_SMS Sent



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?**

Ans- As mentioned in previous answer, and because all 3 are dummy variables in the model are:
1. Tags_Closed by Horizzon: 9.91
2. Tags_Will revert after reading the email: 7.297.
3. Tags_Others: 3.563

These variables have the highest positive coefficients, indicating they have the strongest influence on increasing the probability of lead conversion in the model. Focusing on these three dummy variables can significantly boost lead conversion rates. Specifically, leads tagged as "Closed by Horizzon," "Will revert after reading the email," and " Last Activity_SMS Sent " are more likely to convert.

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.**

Ans- Given the performance of the logistic regression model and the goal of aggressively converting leads during a 2-month intern phase, the company can take following measures:

1. **Leverage High Sensitivity:** Since sensitivity (recall) on the test set is 90.23%, it indicates that the model is effective at identifying most of the actual conversions (true positives). This means that the model can successfully predict leads who are likely to convert. So, use the model's predictions to **prioritize phone calls** to leads predicted as "1" (i.e., potential conversions). With interns available, focus on making personal phone calls to as many of these leads as possible.

2. **Reduce decision Threshold**: Experiment with different thresholds (e.g., 0.3, 0.25) and analyze the precision-recall tradeoff. A lower threshold will result in more leads being predicted as conversions, providing the sales team with more contacts to follow up on.

3.**Prioritize leads** with the highest predicted conversion probabilities.
4. Assign specific segments to each intern (based on lead characteristics) and track their success rate. Identify high-performers and optimize resource allocation accordingly.

This strategy will allow X Education to take full advantage of the interns and ensure they maximize the number of leads converted during this aggressive sales push.

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- Strategy to Minimize Useless Phone Calls:
1. **Increase the Decision Threshold**: This will make the model more conservative, only marking leads with very high probabilities as likely to convert. This threshold will decrease false positives.
2. **Prioritize high-value leads** based on strong positive indicators (like high-impact tags or lead sources).
3. **Focus on Precision,** meaning that when the model predicts a lead as a conversion, it's highly likely to be correct.