Answers of all the questions

Top Three Contributing Variables

Based on the coefficients of the logistic regression model, the top three variables with the strongest positive impact on the probability of a lead getting converted are:

- **Total Time Spent on Website**: The more time a prospect spends on the website, the higher their likelihood of converting. This is an excellent indicator of interest and engagement.

- **Last Notable Activity_SMS Sent**: A lead's last notable activity being an SMS sent by X Education is a strong predictor of conversion. This suggests that the SMS channel is highly effective for converting leads.

- **Tags_Will revert after reading the email**: This tag, which indicates a prospect's positive intent to follow up, is a significant positive predictor. This suggests that the sales team's initial interaction and tagging of the lead are highly correlated with conversion.

2. Top Three Categorical/Dummy Variables to Focus On

The top three categorical variables to focus on for increasing lead conversion are:

- **Tags_Will revert after reading the email**: This is the most impactful tag. The sales team should prioritize leads with this tag and ensure a timely follow-up to capitalize on their high intent.

- **Last Notable Activity_SMS Sent**: The model shows that SMS communication is very effective. The company should consider optimizing its SMS campaigns and increasing their use for lead nurturing.

- **Lead Source_Olark Chat**: Leads originating from Olark Chat are often "hotter" and have a higher propensity to convert. The sales team should be trained to handle these chats effectively and quickly.

3. Strategy for Aggressive Lead Conversion (2-month intern phase)

During this period, the company's goal is to convert as many potential leads as possible. The key is to increase the number of leads classified as "hot" by the model.

Strategy: Lower the classification threshold.

- **Explanation**: The default classification threshold for logistic regression is 0.5. Leads with a predicted probability above 0.5 are classified as "1" (Converted),

and those below are "0" (Not Converted). To be more aggressive, the company can lower this threshold to, for example, 0.3.

- **Impact**: Lowering the threshold will increase the number of leads predicted as "1". This will increase the **recall** of the model, meaning it will correctly identify a higher percentage of the actual converting leads. While this will also increase the number of false positives (leads predicted to convert but do not), the sales team can use the extra manpower from the interns to follow up on these additional leads, ensuring fewer genuine opportunities are missed.

4. Strategy for Minimizing Useless Phone Calls

During periods when the quarterly target is met, the company wants to minimize unnecessary phone calls to save resources and focus on other tasks. The goal is to maximize the efficiency of phone calls.

Strategy: Raise the classification threshold.

- **Explanation**: To ensure only the most promising leads are called, the company should raise the classification threshold to, for example, 0.7 or 0.8.

- **Impact**: Raising the threshold will reduce the number of leads predicted as "1." This will increase the **precision** of the model, meaning that a higher percentage of the leads predicted as "1" will actually convert. This minimizes false positives, ensuring the sales team spends its limited time and resources on the leads most likely to convert, thereby minimizing "useless phone calls."