

# **Lead Scoring Assignment - Summary Report**

## **Lead Scoring Case Study**

Identification of Hot Leads to focus more on them and thus enhancing the conversion ratio.

## **Business Objective**

We are required to categorize or cluster the leads as hot leads and cold leads. This is to be done by analyzing past data provided by X Education Company. The Company's employee will then focus on communicating effectively with the hot leads so that most of them convert. We have been given a target of 80% conversion rate, and thus, to achieve that, we must accurately categorize the leads.

---

## **SUMMARY REPORT**

After conducting extensive exploratory data analysis (EDA), the following key findings can be summarized:

### **1. Univariate Analysis Findings:**

- Page Views Per Visit and Total Visits show low variation, while Total Time Spent on Website has higher variability.
- Page Views Per Visit and Total Visits exhibit a significant number of outliers that require preprocessing before modeling.

### **2. Bivariate Analysis Findings:**

- There is a positive correlation between Total Time Spent on Website and Conversion.
- Conversion shows some correlation with categorical columns like Lead Origin and Lead Source.

### **3. Correlation Analysis (Heatmap) Findings:**

- Total Time Spent on Website has a positive correlation with Conversion.
- Page Views Per Visit and Total Visits have negligible correlation with Conversion.

## **Additional Findings (EDA II):**

### **1. Bivariate Analysis Findings:**

- Certain dummy variables, such as Lead Origin\_Lead Add Form, Last Activity\_Email Bounced, Lead Source\_Reference, and Total Time Spent on Website, have a noticeable impact on Conversion.
2. Correlation Analysis (Heatmap) Findings:
    - Due to the large number of columns, the heatmap lacks clarity. However, significant correlations are observed among the columns, particularly with the dummy variables. It is advisable to further reduce the number of columns to mitigate multicollinearity.

### **Building the Model:**

Two suitable models were built using Logistic Regression with Recursive Feature Elimination (RFE):

1. Model with Moderate Accuracy: This model allows for a wider range of audience targeting, assuming more resources are available for engagement.
2. Model with High Accuracy: This model focuses on a narrower range of potential leads to target, assuming limited resources and the goal of minimizing engagement.

### **Final Model Summary:**

1. Overall accuracy on Test set: 78.6%
2. Sensitivity of the logistic regression model: 73.3%
3. Specificity of the logistic regression model: 82.3%
4. Top 3 variables contributing to lead conversion:
  - Total Time Spent on Website
  - Last Notable Activity\_SMS Sent
  - TotalVisits

### **Recommendations:**

Based on the analysis, X Education Company should consider the following recommendations to improve the lead conversion rate:

1. Focus on Last Activity\_SMS Sent, as it has a positive impact on lead conversion.
2. Address and improve the negative impact of Last Activity\_Olark Chat Conversation.
3. Enhance the Lead Source\_Olark Chat to mitigate its negative impact on conversion.

In conclusion, the Logistic Regression Model shows satisfactory performance compared to the PCA model. To improve the overall conversion rate, X Education Company should prioritize user engagement, increase SMS notifications, boost TotalVisits through advertising, and enhance the Olark Chat service.

### **Additional Recommendations:**

1. Enhancing user engagement on their website: Increased user engagement has a direct positive impact on conversion rates. Therefore, efforts should be made to optimize user experience, encourage active participation, and provide valuable content to keep visitors engaged. This can include interactive features, personalized recommendations, easy navigation, and clear calls-to-action.
2. Increasing the frequency of sending SMS notifications: Leveraging SMS notifications has proven to be an effective strategy in driving higher conversion rates. By implementing targeted and personalized SMS campaigns, X Education Company can effectively engage with potential leads and guide them towards conversion. This can involve sending timely updates, course recommendations, special offers, and reminders to keep the company top-of-mind for potential customers.
3. Enhancing Total visits through advertising initiatives: Boosting the number of website visits, particularly through well-executed advertising campaigns, contributes to higher conversion rates. X Education Company should explore various advertising channels and strategies to attract a larger audience and generate increased interest in their offerings. This can include paid search advertising, display advertising, social media advertising, and partnerships with relevant industry websites or influencers.
4. Improving the Olark Chat service: The current state of the Olark Chat service is negatively impacting the conversion rate. It is crucial for X Education Company to address any shortcomings or issues with the chat service promptly. Enhancements to its functionality, responsiveness, and user experience can help foster better communication with potential leads and ultimately improve conversion rates. This may involve training chat operators to provide timely and accurate information, implementing proactive chat invitations, and integrating chat data with lead management systems for seamless follow-up.

By implementing these recommendations, X Education Company can optimize its lead conversion process and increase the probability of converting potential leads into paying customers. These strategies aim to improve user engagement, leverage effective communication channels, drive website traffic, and enhance the chat service, ultimately contributing to higher conversion rates and overall business success.

---

### **Team Members:**

Piyush Gupta | Prashant Singh | Arun Tiwari