Lead Scoring Analysis for X Education: Summary Report

Introduction

X Education, an online course provider, faces a challenge with its lead conversion rate, which stands at approximately 30%. Despite acquiring numerous leads through various marketing channels, a majority do not convert into paying customers. To enhance the efficiency of the sales team and improve conversion rates, we implemented a **logistic regression-based lead scoring model** to identify high-potential leads, also known as "Hot Leads."

Data and Methodology

The dataset used for this analysis consists of **9,240 records** and **37 attributes**, including features such as Lead Source, Last Activity, Total Time Spent on Website, and various categorical tags. The target variable, **Converted (1/0),** indicates whether a lead was successfully converted or not.

The key steps in the analysis were:

1. **Data Preprocessing:**
   o Handled missing values by imputing categorical modes and median for numerical fields.
   o Converted categorical variables into dummy variables.
   o Standardized numerical features to ensure consistency in model performance.
2. **Feature Selection:**
   o Used **Recursive Feature Elimination (RFE)** to select the most influential predictors.
   o Identified the top 10 contributing variables.
3. **Model Building:**
   o Applied **Logistic Regression** to assign lead scores.
   o Evaluated model performance using **accuracy, confusion matrix, and classification report.**

Key Findings

The analysis revealed the **top three predictors** of lead conversion:

1. **Lead Source - Welingak Website**: Leads from this source have a higher probability of conversion.
2. **Last Activity - SMS Sent**: Engaging with leads via SMS significantly influences conversion rates.
3. **Tags - Closed by Horizzon**: This indicates successful closure of leads by specific processes.

Other notable factors include tags related to "Ringing," "Will revert after reading the email," and "Wrong number given," highlighting key patterns in lead behavior.

Business Recommendations

Based on our findings, we propose the following strategies for X Education:

1. **Aggressive Conversion Strategy (Intern Period):**
   - Interns should **prioritize high-scoring leads**, particularly those with "Ringing" or "Will revert" tags.
   - Increase follow-up efforts via SMS and emails before making calls.
   - Use a multi-channel approach to engage leads actively.
2. **Minimal Calls Strategy (When Targets Are Met):**
   - Focus only on high-probability leads.
   - Avoid calling leads with "Wrong number given" or "Switched off" tags to minimize wasted efforts.
   - Implement email nurturing campaigns instead of direct calls for low-priority leads.

Conclusion

By leveraging logistic regression for lead scoring, X Education can **enhance sales efficiency, prioritize high-potential leads, and reduce wasted efforts**. This data-driven approach will ultimately help improve the overall conversion rate and maximize revenue generation.