Lead Installation Forecasting for Broadband Provider

BACKGROUND

A leading broadband provider in India is analyzing its lead generation and conversion processes across multiple cities. Despite generating thousands of leads monthly through **social media**, **website interactions**, **referrals**, and **walk-ins**, the company faces challenges with:

- **High customer acquisition costs**
- **Inefficient resource allocation**

Although the company has a robust lead-generation strategy, it struggles to identify leads with the highest likelihood of conversion into successful installations. The company aims to optimize its marketing efforts and allocate resources effectively to pursue high-value leads. Leveraging **business analytics** is critical to uncover insights that streamline operations, refine lead management, and improve conversion rates.

BUSINESS PROBLEM

1. Optimization of Lead Funnel

Using the data, propose business strategies to optimize the lead conversion process, focusing on:

- **Customer Segmentation:** How can lead segmentation improve to ensure resources target high-conversion potential leads?
- **Marketing Spend Optimization:** How should the marketing budget be allocated across channels to maximize ROI and reduce customer acquisition costs?
- **Lead Follow-Up Prioritization:** How should follow-up efforts be optimized to ensure high-potential leads are pursued first?

2. Predict Lead Conversion

Develop a predictive model to identify leads more likely to result in successful installations based on provided features.

- Focus areas: Marketing channels, lead engagement, and operational efficiency.
- **Model Objective:** Prioritize high-likelihood leads, reduce wasted efforts, and improve overall resource efficiency.

DATASET OVERVIEW

The dataset contains **100,000 rows** and features spanning customer behavior, financial performance, and operational metrics. Key features include:

- **Customer Information:** Lead source, plan type, income level, city, tech-savviness.
- **Customer Engagement Metrics:** Number of follow-ups, competitor interest, preferred contact time.
- **Operational Metrics:** Technician visits, service resolution time, days to acceptance.
- **Financial Data:** Marketing spend, discount availed, lifetime value, plan cost.
- **Funnel Progression Metrics:** Days to accept, qualify, install request, final installation status.
- **Seasonality & Temporal Features:** Month of creation, day of the week, holiday status, festive periods.

Target Variable:

- **Installed:** Indicates whether a lead was converted into a successful installation.

DATA DICTIONARY

Column Name	**Description**
Lead ID	Unique identifier for each lead.
City	The city in which the lead was generated.
Plan Type	Type of broadband plan chosen by the lead.
Household Income Level	The income level of the lead’s household.
Lead Source	The source from which the lead originated.

Marketing Channel	The marketing channel through which the lead was acquired.
Lead Created Date	The date the lead was created.
Follow-Up Count	The number of follow-up attempts made on the lead.
Competitor Interest	Whether the lead showed interest in competitor services.
Preferred Contact Time	The lead's preferred time for contact.
Customer Tech-Savviness	The technology savviness level of the customer.
Decision Influence	The main factor influencing the customer's decision to purchase.
Complaint History	Whether the lead has a complaint history related to previous services or products.
Payment Mode Preferred	The payment mode preferred by the lead.
Payment Frequency	The preferred frequency for payments.
Discount Availed (INR)	The discount amount availed by the lead, in Indian Rupees.
Marketing Spend (INR)	The marketing spend attributed to acquiring this lead, in Indian Rupees.
Plan Cost (INR)	The cost of the broadband plan selected by the lead, in Indian Rupees.
Average Monthly Spend (INR)	Average monthly spend by the lead.
Lifetime Value (INR)	The lifetime value (LTV) of the customer.
Bundled Service Interest	Whether the lead showed interest in bundled services.
Competitor Price Sensitivity	Sensitivity of the lead to competitor pricing.
Distance to Service Hub	The distance (in kilometers) from the lead's location to the nearest service hub.

Infrastructure Ready	Whether the necessary infrastructure is available at the lead's location for installation.
Network Downtime (Hours)	The number of hours the lead has experienced network downtime in the past month.
Signal Strength	The strength of the broadband signal available at the lead's location.
Service Quality Rating	A rating of the broadband service quality (1 to 5).
Time Spent on Research (Days)	The number of days the lead spent researching broadband options before making a decision.
Preferred Communication Mode	The preferred method of communication for the lead.
Days to Accept	The number of days between lead creation and acceptance.
Days to Qualify	The number of days between lead acceptance and qualification.
Days to Install Request	The number of days between lead creation and installation request.
Installed	Whether the lead was successfully installed (1 = Installed, 0 = Not Installed).
Is Holiday	Whether the lead was created during a public holiday.
Festive Period	Whether the lead was created during a festive period (October-December).
Day of the Week	The day of the week the lead was created.

DELIVERABLES

1. Data Preprocessing & EDA Summary

- A report on data cleaning, initial exploratory data analysis (EDA), and key insights.

- Highlight issues like **missing values**, **outliers**, and correlations between features and the target variable.

2. Feature Importance Insights

- Identify key features impacting successful installations and their business relevance.
- Suggest factors the company should prioritize in marketing and operations.

3. Strategic Recommendations

Based on model insights, provide actionable strategies to:

- Improve **lead prioritization** and **resource allocation**.
- Optimize **marketing channels** and **customer engagement** strategies.
- Enhance operational processes to increase conversion rates.

4. Analysis Documentation

Submit the analysis file documenting the entire process, from data preprocessing to model building and evaluation, in one of the following formats:

- **Python notebook (HTML format)**
- **Excel sheet (XLSX)**
  Include key findings to enable data-driven business decisions.

5. Model Performance Summary

- A summary comparing model performance metrics, emphasizing model explainability and business applications.

RULES FOR SUBMISSION

1. 3-Slider Deck

- Submit a **3-slide PDF presentation** (excluding title and thank-you slides).
- Ensure the slides are concise, clear, and aligned with the business problem.

2. Analysis File

Submit the analysis file in one of the specified formats:

- **Python notebook (HTML format)**
- **Excel sheet (XLSX format)**

3. Submission Deadline & Format

- Late submissions or those in formats other than specified will not be considered. The decision of the organizers is final and binding.

Contact for Queries

- **Shikhar Mishra:** +91 9981195566
- **Srishilesh P S:** +91 7598480495