Insurance Cross-Selling Project – Guidelines

Project Overview

AssurePredict is a leading insurance company specializing in innovative risk management solutions. This project aims to develop a predictive model to identify potential cross-selling opportunities among existing clients, specifically targeting those who might be interested in purchasing an additional vehicle insurance policy.

Project Objective

The goal is to build a **machine learning model** that predicts whether clients who currently have health insurance might be interested in subscribing to a vehicle insurance policy. The model will help AssurePredict improve the effectiveness of its cross-selling strategies and increase market penetration.

Value for AssurePredict:

- Increased conversion rates in auto insurance sales.
- Optimized marketing campaigns by targeting clients most likely to purchase.
- Reduced costs from ineffective campaigns through precise targeting.

Dataset

The dataset is available for download here and contains detailed information about clients and their insurance behavior.

Key features:

- id: Unique client identifier.
- Gender: Client gender.
- Age: Client age.
- Driving_License: 1 if the client holds a driving license, 0 otherwise.
- Region_Code: Unique code of the client's region of residence.
- Previously_Insured: 1 if the client already has vehicle insurance, 0 otherwise.
- Vehicle_Age: Age of the client's vehicle.
- Vehicle_Damage: 1 if the client has experienced vehicle damage or accidents, 0 otherwise.
- Annual_Premium: Annual insurance premium paid by the client.
- PolicySalesChannel: Channel used to sell the policy (e.g., email, phone, in person).
- Vintage: Number of days the client has been insured with AssurePredict.
- Response: 1 if the client accepted the cross-sell offer, 0 otherwise.

Required Activities

1. Dataset Exploration

Perform an initial exploration to understand the distribution of features and the target variable. Focus on:

- Distribution of the Response variable to identify potential class imbalance.
- Relationships between key variables such as Annual_Premium, Vehicle_Age, Previously_Insured, and client response.

Value: Thorough data exploration helps uncover hidden patterns and critical points that will influence the success of the predictive model.

2. Handling Class Imbalance

The target variable Response may be imbalanced, with many more clients declining the cross-sell offer than accepting it. Address this issue using techniques such as:

- **Class Weights:** Penalize the majority class in the model.
- **Oversampling or Undersampling:** Create a more balanced dataset to improve model generalization.

Value: Properly managing class imbalance is crucial to avoid high false-negative rates and improve the accuracy of cross-selling predictions.

3. Predictive Model Development

Build a machine learning model to predict the probability that a client will respond positively to the cross-sell offer.

Value: The predictive model allows AssurePredict to accurately identify clients most likely to purchase an additional policy, improving marketing ROI.

Conclusion

This project enables AssurePredict to leverage machine learning for effectively identifying cross-selling opportunities. Adopting a **data-driven approach** ensures not only increased sales but also higher customer satisfaction through more relevant and personalized offers.