Lead Scoring Case Study - Summary Report

Introduction

X Education aims to enhance its lead conversion process by utilising a structured analytical approach. The objective of this study was to develop a predictive model that assigns a lead score between 0 and 100, helping the sales team focus on high-potential leads. The ultimate goal was to improve the conversion rate from 38% to an optimal 80%.

Methodology

The study followed a structured process:

1. **Data Collection & Inspection** - Examined the dataset for completeness and consistency.
2. **Data Cleaning & Preparation** - Removed irrelevant variables, handled missing values, and treated outliers.
3. **Exploratory Data Analysis (EDA)** - Analysed categorical and numerical features to identify key trends.
4. **Feature Engineering** - Created dummy variables and normalised numerical variables.
5. **Train-Test Split** - Divided the dataset into 70% training and 30% testing.
6. **Feature Selection** - Used Recursive Feature Elimination (RFE), Variance Inflation Factor (VIF) analysis, and correlation checks to retain relevant features.
7. **Model Building** - Trained multiple logistic regression models, iteratively refining based on performance metrics.
8. **Model Evaluation** - Assessed the final model's accuracy, precision, recall, and ROC-AUC score.
9. **Business Insights & Strategy Development** - Provided actionable recommendations based on model findings.

Key Findings & Results

1. Top Features Affecting Lead Conversion

- **Total Time Spent on Website**: Leads spending more time on the site had higher conversion rates.
- **Total Visits**: A higher number of visits indicated greater engagement and likelihood of conversion.
- **Lead Source**: Certain sources like referrals and direct visits were more likely to convert.

2. Important Categorical Features

- **Lead Source_Welingak Website** and **Lead Source_Reference**: These sources had high conversion rates.
- **What is your current Occupation_Working Professional**: Working professionals showed a higher probability of conversion.

Business Strategy Recommendations

1. Enhancing Lead Conversion During Intern Hiring Phase

- Target high-intent leads based on website engagement and prior interactions.
- Lower the decision threshold to increase recall and reach more potential leads.
- Monitor conversion performance and adjust strategy accordingly.

2. Reducing Unnecessary Calls When Targets Are Met

- Shift to high precision targeting to avoid low-quality leads.
- Focus follow-ups on leads who previously showed interest.
- Avoid contacting leads with a high likelihood of rejection, such as students or those marked "Do Not Call."

Final Model Results

- **Overall Model Accuracy**: 81%
- **ROC-AUC Score**: 0.89 (indicating strong predictive capability)
- **Optimal Cut-off Probability**: 0.34 (balancing recall and precision)
- **Training Data Performance**:
    - Sensitivity: 81.7%
    - Specificity: 80.6%
- **Test Data Performance**:
    - Sensitivity: 80.4%
    - Specificity: 80.5%

Conclusion & Next Steps

By leveraging logistic regression and data-driven insights, X Education can significantly enhance its lead conversion strategy. The predictive model provides a structured approach to prioritising leads, optimising sales efforts, and aligning with business objectives. Implementing the recommended strategies should improve conversion efficiency and maximise return on investment.