

Summary Report: Lead Scoring Case Study

Problem Statement

An education company named X Education sells online courses to industry professionals. The company identifies potential leads from their website and seeks to improve their lead conversion rate. The goal is to identify 'Hot Leads' to focus sales efforts and increase conversion rates from 30% to 80%.

Business Objective

The company aims to build a model that assigns a lead score (0-100) to each lead. This will help the sales team prioritize leads with higher conversion potential.

Problem Approach

The approach involves data importing, preparation, EDA, dummy variable creation, train-test split, feature scaling, model building (RFE, R-squared, VIF, p-values), and model evaluation.

EDA Insights

Key insights include: 1. Leads from HR, Finance & Marketing specializations have a high probability of conversion. 2. Leads from Google & direct traffic sources are more likely to convert. 3. Unemployed individuals show higher interest in joining the course.

Model Evaluation

The model's performance is evaluated using ROC curve and tradeoff between precision and recall. A prospect lead with a conversion probability above 42% is considered a hot lead.

Final Model Performance

The model achieves 80% accuracy, 77% sensitivity, and 80% specificity on both train and test data.

Conclusions

1. Focus on leads originating from API and Landing page submissions. 2. Maximize efforts on Google/direct traffic leads and reference conversions. 3. Leads spending more time on the website are more likely to convert. 4. Most conversions occur with unemployed and working professionals.