

In our quest to bolster X Education's appeal and attract more industry professionals to their courses, we delved deep into the data trenches. Here's the breakdown of our findings:

1. **Data Cleanup** : We embarked on a tidying mission, addressing null values and substituting those ambiguous 'select' options with nulls. Additionally, we categorized some nulls as 'not provided' to minimize data loss. Furthermore, we streamlined location information to 'India', 'Outside India', and 'not provided'.
2. **Exploratory Data Analysis (EDA)** : A swift examination revealed that certain categories weren't contributing much, while numerical data appeared robust with no outliers skewing the results.
3. **Dummy Variables** : With 'not provided' entries out of the picture, we forged ahead and generated dummy variables. Numeric data underwent scaling using MinMaxScaler to level the playing field.
4. **Data Splitting** : To ensure a fair assessment, we partitioned the data into 70% for training and 30% for testing purposes.
5. **Model Development** : Employing Recursive Feature Elimination (RFE), we pinpointed the top 15 contenders. Subsequently, through manual intervention guided by Variance Inflation Factor (VIF) and p-values, we curated the final lineup, retaining variables with  $VIF < 5$  and  $p\text{-value} < 0.05$ .
6. **Evaluation Metrics** : Armed with a confusion matrix, we embarked on a quest for the optimal cutoff value, employing the mystical powers of ROC curves. Our endeavors culminated in achieving commendable accuracy, sensitivity, and specificity rates, all hovering around the 80% mark.
7. **Predictive Analysis** : Venturing into the realm of prophecy, we unleashed our predictive prowess upon the test data, uncovering an optimal cutoff threshold of 0.35, accompanied by the same impressive 80% scores.

In essence, the heavyweight contenders influencing potential buyer behavior include:

1. Total time spent on the website.
2. Total number of visits.
3. Lead sources:
  - Google

- Direct traffic
- Organic search

4. Last activities:

- SMS
- Olark chat conversation

5. Lead origin as Lead add format.

6. Current occupation as a working professional.

Armed with these invaluable insights, X Education stands poised to capture the attention of nearly all prospective buyers and compel them to take the plunge into their courses.