For X Education, an analysis was conducted to determine which industry professionals are most likely to enroll in their courses. The dataset provided includes information on customers who visit the website, the duration of their visit, how they arrived at the website, and the conversion rate. The following technical approaches were used:

1. Data Cleaning: Redundant variables were removed, 'Select' was replaced with a null value, null values with a percentage higher than 40 were dropped, missing values were imputed with the max related value, highly skewed columns were identified and dropped, missing values were treated by imputing a favorable aggregate function (Mean, Median, Mode), and outliers were detected.
2. Exploratory Data Analysis: EDA was performed on the data and it was found that many factors in the categorical variables were insignificant. Numerical values appeared satisfactory but outliers were found. Univariate analysis was conducted on both categorical and continuous variables and bivariate analysis was performed with respect to the target variable.
3. Dummy Variables: Dummy variables were created for each categorical column.
4. Scaling: Data for continuous variables was scaled using a Standard scalar.
5. Train-Test Split: The split for train and test data was performed at 70% and 30%, respectively.
6. Model Building: RFE was used to list the most important variables. Variables with a VIF of 5 and a p-value of 0.05 were kept while irrelevant features were manually removed based on VIF values and p-value.
7. Model Evaluation: The accuracy, sensitivity, and specificity scores were fair at 81%, 79%, and 83% for the test set and 81%, 80%, 81% for the train set respectively.
8. Conclusion: The total time spent on the website, the number of visits, the primary source being Olark Chat, and the last action being an SMS Olark chat conversation are all important factors in determining potential buyers.