

LEAD SCORING CASE STUDY

SUMMARY

Problem Statement

An education company named X Education sells online courses to industry professionals. On a day to day basis the potential learners visit the website and browse for the courses of their interest.

There are other sources like referrals which also contributes to the potential learners visiting the website. Once the potential learners sign-up into the website they are classified as leads. Through the details provided during the time of sign-up the sales team of the X education connect with the leads to get them converted.

The problem statement being minimal number of converters, typically 30 percent. X education want to establish a process which helps to identify most potential leads, known as 'Hot Leads' and work towards better conversion. The expected conversion rate is 80 percent.

Data

Leads dataset from the past with around 9000 data points has been provided. This dataset consists of various attributes such as Lead Source, Total Time Spent on Website, Total Visits, Last Activity, etc. that would help in identifying Hot Leads.

Process Flow

- Data was initially read and understood
- The columns were renamed with meaningful names
- Data was first assed to understand the data types and null values.
- Post the cleaning of null values the redundant columns were worked upon and dropped
- Data was analysed through means of graphs , heat maps and plots
- Data was converted to Binary dataset
- Dummy variables were created
- Outliers were treated
- Data was divided into test and train and scaling was done
- Model was built and assessed

Approach

The major variables used for assessing the hot leads are :

- **Total Website Visit** : The more number of times a lead visits the website, the more likely the lead is to get converted
- **Time spent on website** : The more time spent on assessing the courses provided by x education, the more are the possibilities of the lead opting for a course
- **Lead source** : Helps in understanding the potential platform for advertising , which may attract more leads and hot leads

- **Specialization:** Helps in understanding the courses that have high market preferences
- **Occupation:** Helps in analysing the market population that can be targeted for better conversion

The various combination of the above mentioned variables can be used to identify the potential converters.

Findings

- The highest number of lead origin is from Landing page submission
- The highest lead source is Google
- Most preferred specialization is Management Specialization
- Most population looking for upskilling is unemployed population
- Unemployed leads spending more time visiting the website, could be a potential hot lead as they might be looking towards improving their marketability and could be targeted as their conversion would be easier.
- The increased website visit implies the interest of the lead towards a particular course
- As a specialization Management course have gained high interest among the leads. Focusing on marketing the management courses can help in better conversion
- Mumbai has the largest population of leads. Marketing the course benefits can help in increased interest of the leads, leading to better conversions.