# Leading Score

Problem Statement

- X Education sells online courses to industry professionals. X Education needs help in selecting the most promising leads, i.e. the leads that are most likely to convert into paying customers. The company needs a model wherein you a lead score is assigned to each of the leads such that the customers with higher lead score have a higher conversion chance and the customers with lower lead score have a lower conversion chance. The CEO, in particular, has given a huge target lead conversion rate of around 80%

Business Goal

X Education needs help in selecting the most promising leads. Higher leading score higher conversion rate
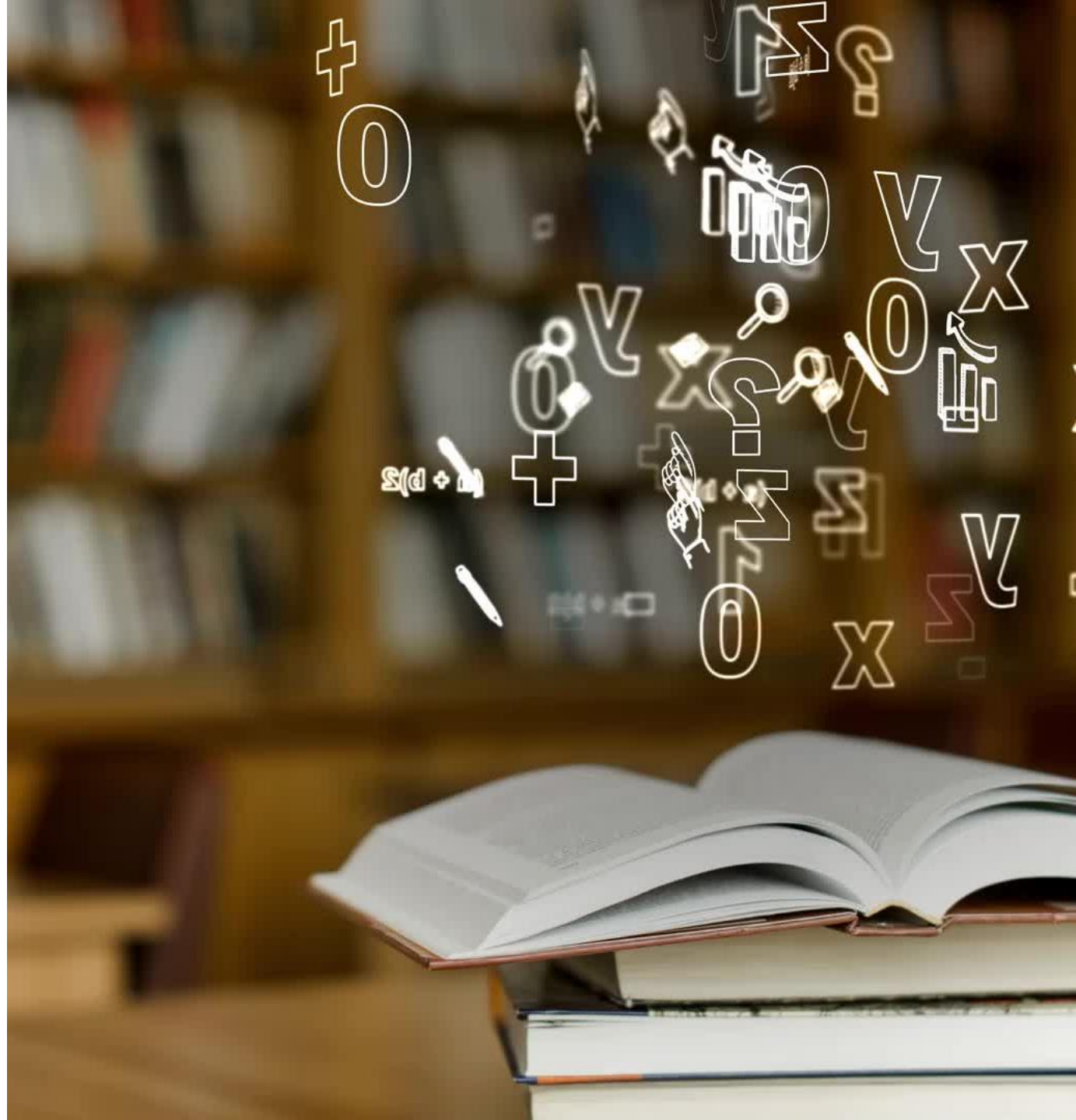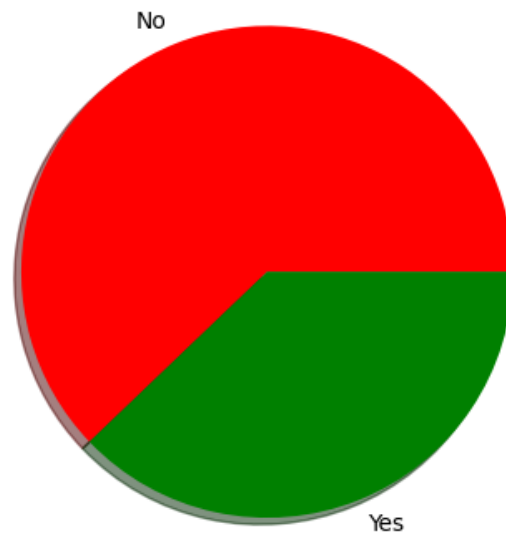
Action

- Import and Inspect Data

- EDA

- Data Preparation

- Train Test Split

- Feature Scaling

- Model Building

- Feature Selection using RFE

- ROC curve

- Optimal Cutoff Point

- Making prediction on the test set

- Assigning a Lead Score based on the

- Model

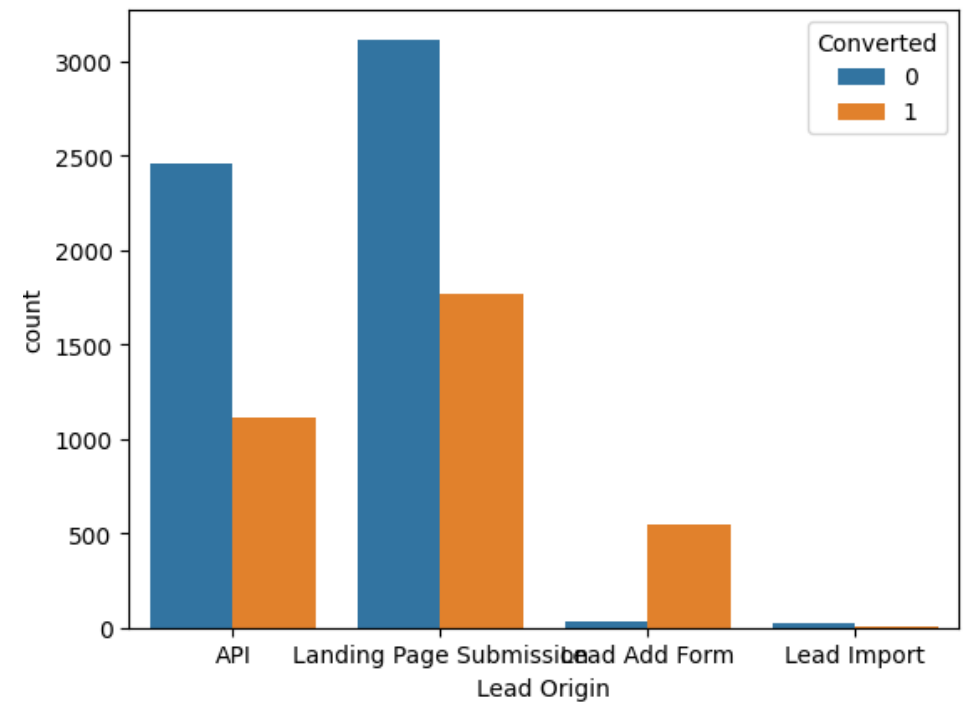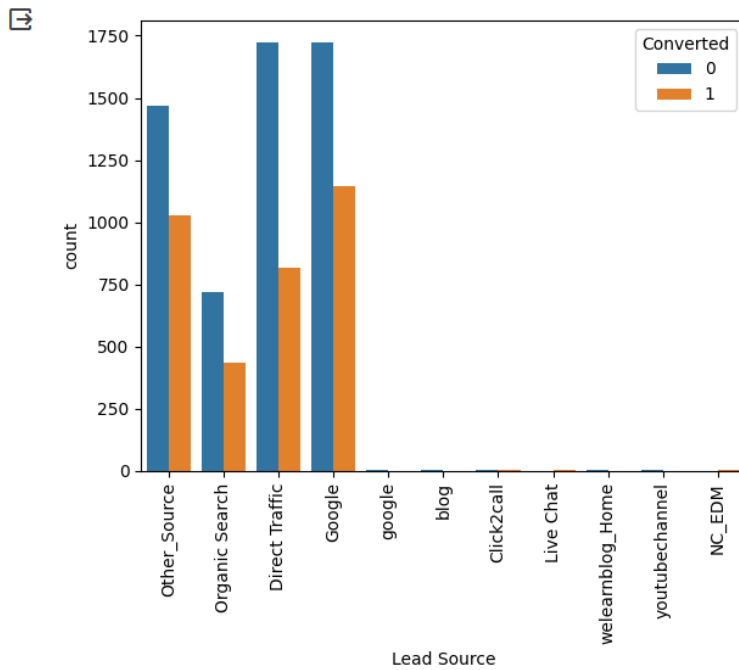- Feature importance of the final model
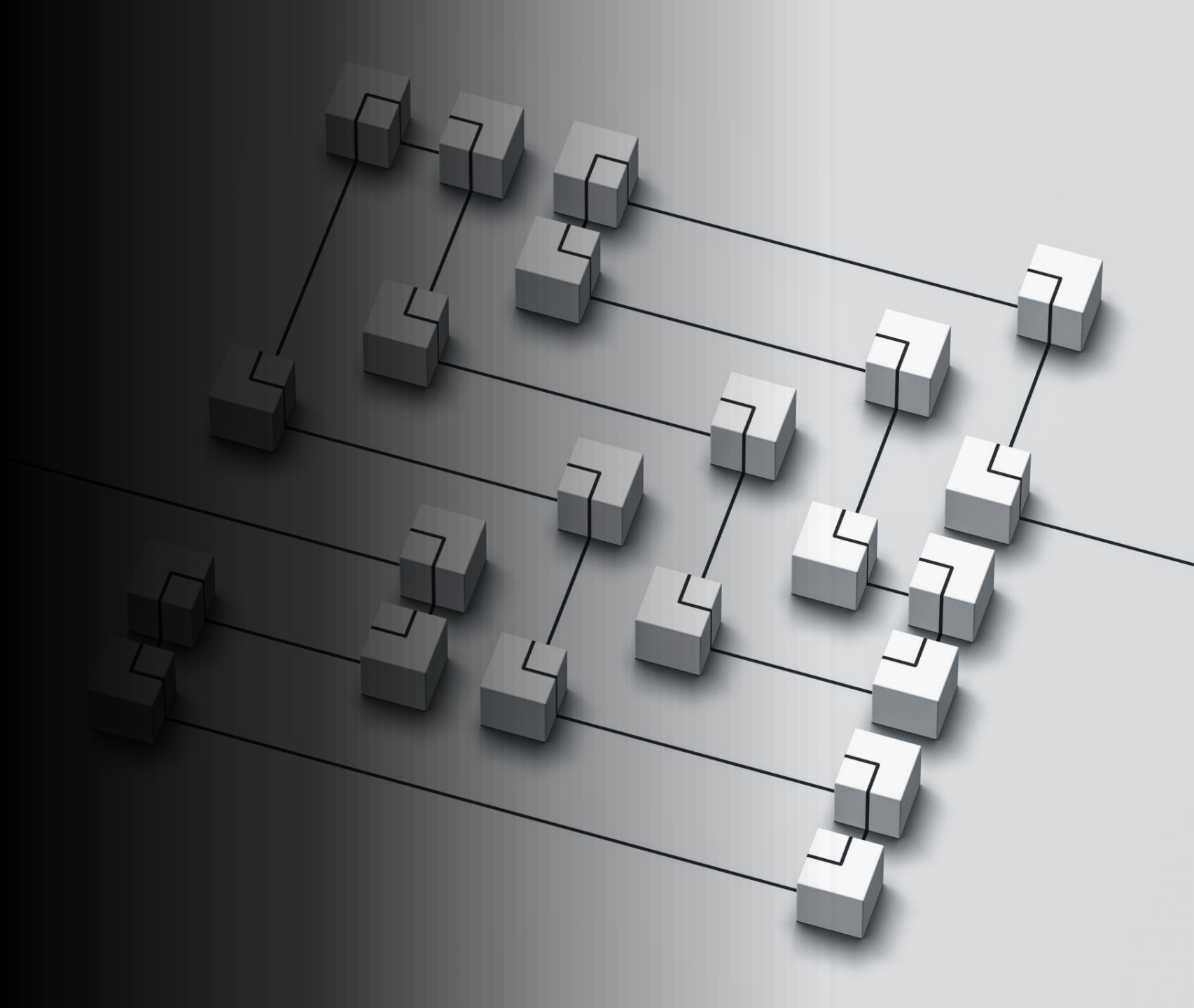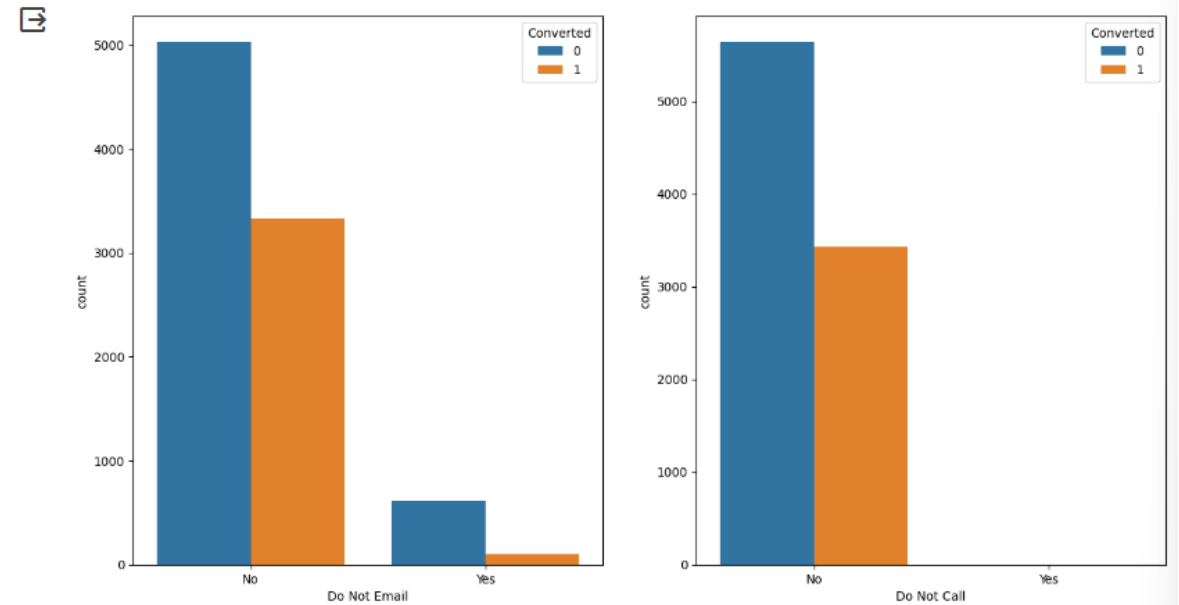
EDA

- Converted

It is not a hugely imbalanced data

EDA

1. Direct Traffic and Google generates maximum number of Leads
2. Conversion rates of all other sources combined is higher but they are in very small numbers
3. The conversion rates of direct traffic needs to be improved using strategies like promotional discounts on direct traffic

1. People who opt for Do not Email, mostly does not convert since it is already implied that they do not want to be contacted further.
2. People who does not opt for any of these DND choices have almost equal chance of converting and not converting.

Importance features

```
Total Time Spent on Website                              100.000000
Lead Origin_Lead Add Form                                 64.832962
What is your current occupation_Working Professional      58.721321
Last Activity_Other Activity                              48.666831
Last Activity_SMS Sent                                    30.077787
Last Activity_Unsubscribed                                29.662509
What is your current occupation_Other Occupation          22.798279
Lead Source_Other_Source                                  21.935381
Last Activity_Unreachable                                 17.378989
Last Activity_Olark Chat Conversation                    -20.139830
Last Notable Activity_Modified                           -20.815764
Specialization_Other Specialization                      -24.580761
Lead Origin_Landing Page Submission                      -25.062014
Do Not Email                                             -36.721736
```