Data Science use case

Why do a case?

We want to simulate real working conditions for you as an analytics contributor at Nawy. We believe in giving time and space so that you can grasp the problem, explore the data and weave a story that clearly brings out the recommendations and next steps. Following in the footsteps of Einstein who said –“If I had an hour to solve a problem I'd spend 55 minutes thinking about the problem and five minutes thinking about solutions.”

Come back with a story!

What do we expect from the story?

1. A Clean Python Notebook showing your work during the data exploration and feature engineering phase.

2. Document and mention all experiments you did whether in jupyter notebooks or excel sheets.

3. You should have clear next steps and recommendations based on your analysis.

4. Powerpoint presentation showing insights you have found in the data and the best experiment in terms of data,features,model and results

5. Extra points for- modularized and reusable code and model deployment.

6. Delivery will be a github repo including all deliverables.

What is the problem context?

In Nawy we are initiating targeting campaigns on a daily basis to generate Real estate leads. The main pain is generating qualified leads.So we need to build a model to classify the received leads into two categories: low qualified and high qualified leads.

Data Dictionary

Column Name	Description
lead_id	Is unique id representing customer (customer my reach more than one time)
customer_name	Name of the customer
lead_mobile_network	To differentiate between local and international customers
message	Message left by the customer
lead_time	Time of receiving the lead
method_of_contact	Contact method used to reach us
ad_group	Include information of target audience used in the campaign
lead_source	Channel used to reach us
campaign	Name of the targeting campaign
Location	The location which the customer is searching for
Low_qualified	Define if the lead is low or high qualified