**Lead Scoring**

**Understanding the problem statement**

A lead is someone who has shown an interest in a particular product and is very likely to buy the product.

Lead is captured in 2 ways: Direct and Indirect ways.

Direct ones are captured by the users filling out the form. Indirect ones are captured by the users who have not filled out the form but visited the page. You are provided with the direct leads who filled out the forms.

Lead is captured from the forms. Any user who fills out the form is known as a lead. Identifying the potential leads beforehand will help us to prioritize the leads and pitch our product to these users first.

Potential leads have a higher chance of buying the product compared to other leads. We can identify the potential leads by assigning every lead a score.

**Objective**

The objective of the problem is to build a smart system to assign a score to the user based on his user journey.

**Features**

Created\_at: date and time at which lead was dropped

Subdomain: domain of the email id

Ext: extension of the email id

Country\_code: country code of the user

Form\_type: type of the form through which the user has dropped a lead

Experience: experience of the user

Utm\_source: UTM source of the lead

Utm\_medium: UTM medium of the lead

Is\_signedup: boolean flag indicating if the user has signed up while dropping the lead

vintage: vintage of the user

DateTime features:

weekday

quarter

is\_weekend

is\_month\_start

is\_month\_end

is\_quarter\_start

is\_quarter\_end

#leads: no. of leads dropped in the last 3 months

lead\_rate: no. of leads/time

Form\_{1-15}: no. of leads dropped through different forms

buy: target variable indicating if the user has bought the product in the next 90 days after dropping the lead