**Insurance Premium Default Propensity**

**Objective**: Premium paid by the customer is the major revenue source for insurance companies. Default in premium payments results in significant revenue losses and hence insurance companies would like to know upfront which type of customers would default premium payments.   
The objective of this project is to predict the probability that a customer will default the premium payment, so that the insurance agent can proactively reach out to the policy holder to follow up for the payment of premium.

**About the dataset**

The dataset contains the following information about 79854 policy holders:

1. id: Unique customer ID
2. perc\_premium\_paid\_by\_cash\_credit: What % of the premium was paid by cash payments?
3. age\_in\_days: age of the customer in days
4. Income: Income of the customer
5. Marital Status: Married/Unmarried, Married (1), unmarried (0)
6. Veh\_owned: Number of vehicles owned (1-3)
7. Count\_3-6\_months\_late: Number of times premium was paid 3-6 months late
8. Count\_6-12\_months\_late: Number of times premium was paid 6-12 months late
9. Count\_more\_than\_12\_months\_late: Number of times premium was paid more than 12 months late
10. Risk\_score: Risk score of customer (similar to credit score)
11. No\_of\_dep: Number of dependents in the family of the customer (1-4)
12. Accommodation: Owned (1), Rented (0)
13. no\_of\_premiums\_paid: Number of premiums paid till date
14. sourcing\_channel: Channel through which customer was sourced
15. residence\_area\_type: Residence type of the customer
16. premium : Total premium amount paid till now
17. default: Y variable - 0 indicates that customer has defaulted the premium and 1 indicates that customer has not defaulted the premium