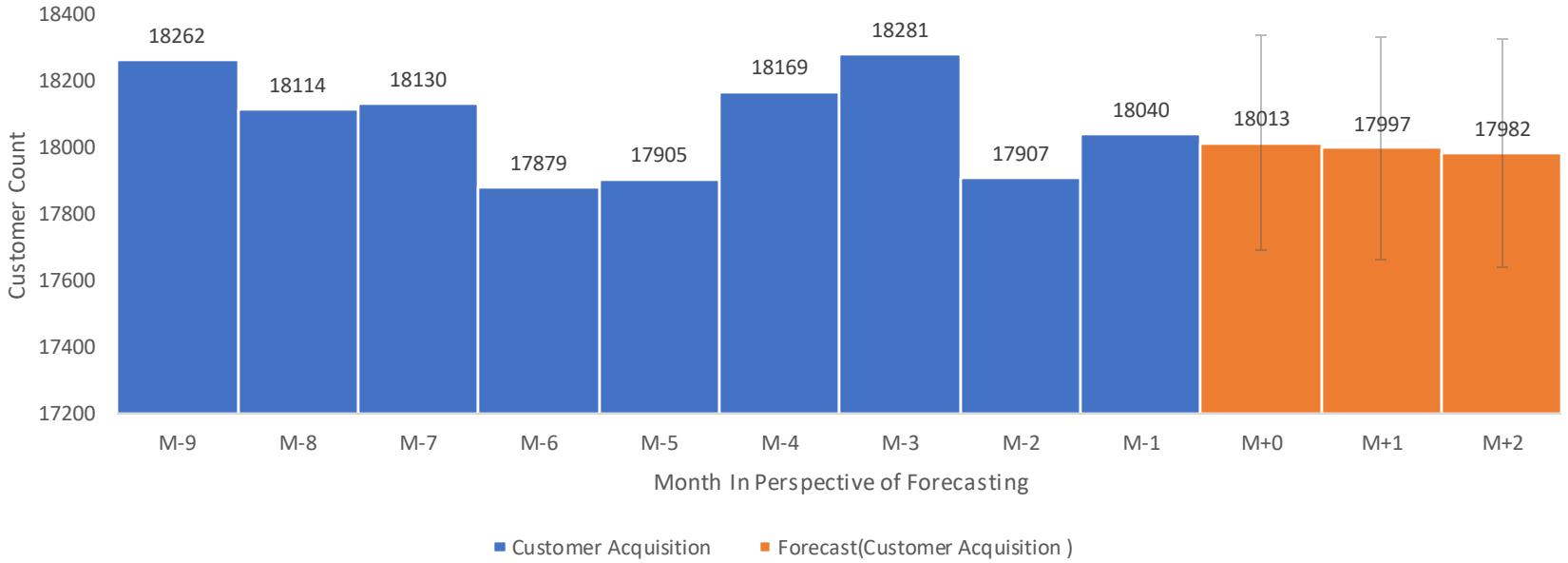
Prodyogiki - An IT and Analytics Based Case Study Competition

Team Vishleshak
Roshan Kumar Biswal
Parul Wategaonkar
Utkarsh Arora

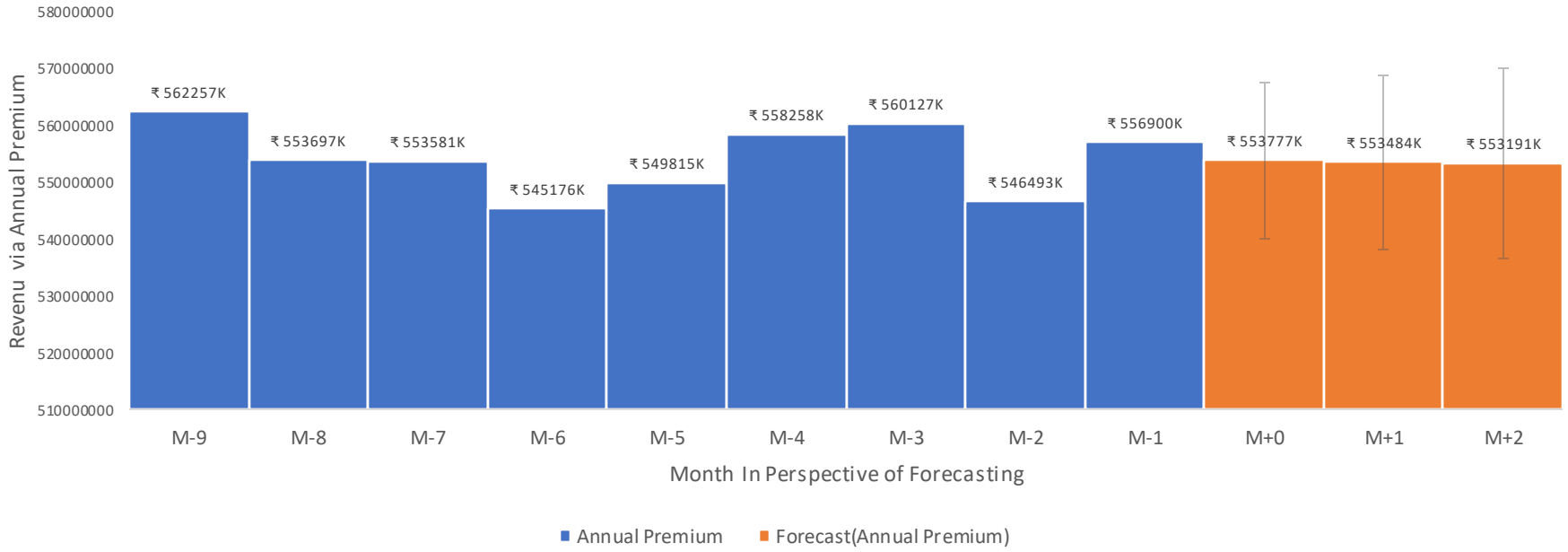
Predictive Modeling To Forecast Future Sales

- Forecasting the **Number of Customer** and **Revenue Generated** based on **Customer Association** only for customers who have taken the insurance (insured = 1)
- Using Customer Association dates a Customer is associated to M-XX (if a customer is associated for 30 days; he is assigned to M-1).**Customer Association date < 271** is used to forecast **M-0,M+1,M+2**). Assumptions: Considering the data for 270 days we have divided it into bins of 30 days each. Thus 9 months data has been used for the projection. The past twenty days i.e. data points from 270 to 290 haven't been included in forecasting as this would have had the least significance on future trends.
- The m-o-m growth rate obtained has been close to -0.14% which clearly points to the stagnant (slightly declining) customers being acquired by the company. This is a major issue of concern for the company.



Predictive Modeling To Forecast Future Revenue

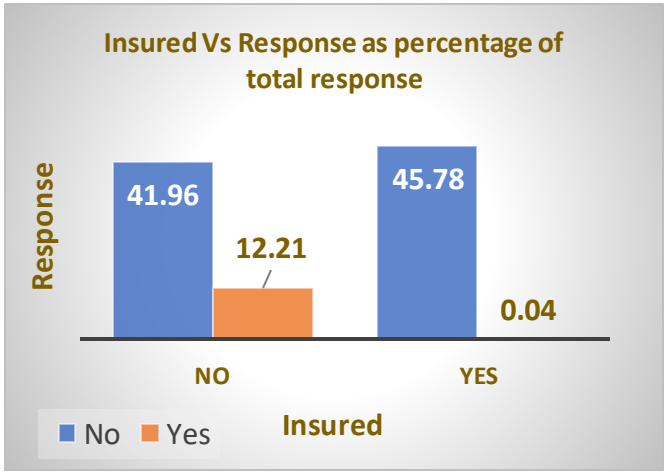
- The forecasting for the revenues has been done similar to number of customers acquired. Thus 9 months duration have been used for the future projection of monthly premiums.
- There is no **seasonality** present in the data which is the reason for no deviation in the trendline. The projected revenues for the upcoming 3 months (**M+0, M+1, M+2**) is almost stagnant. The m-o-m growth rate has been slight negative at around **–0.11%.**
- On an average the revenue generated per customer is around **Rs. 30,649.67.** The number of customers associated needs to be increased in order to increase the total revenue generation of the company.



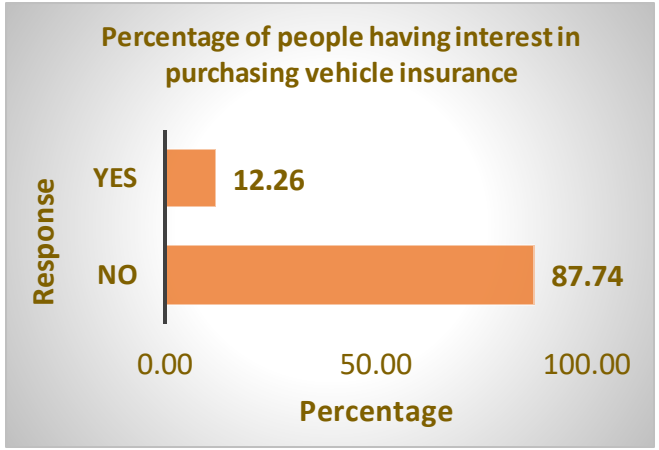
Exploratory Data Analysis

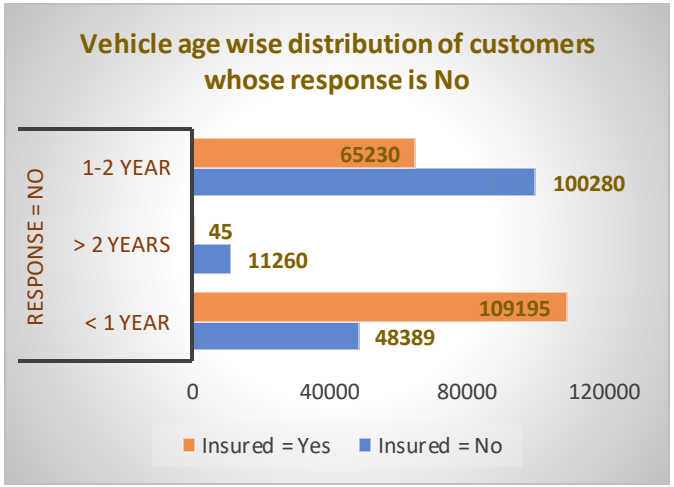
Insured Vs Response

	Response		
		No	Yes
Insured	No	159929	46552
	Yes	174470	158



It has also been observed that only 0.090% of the people having past history of purchasing vehicle insurance has shown interest in purchasing further which is a quite low number

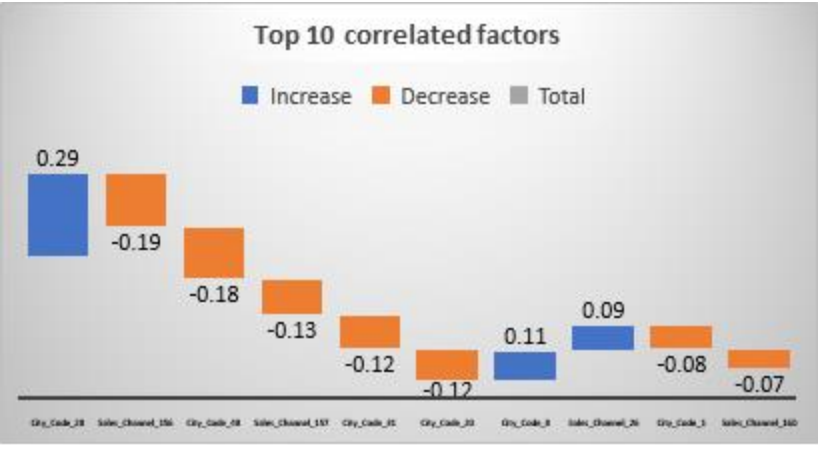




Insights:

- As the number of people having vehicle insurance in past and not interested in purchasing is high as 174470 i.e. 99.90% of the people having past purchasing history, it can be inferred that people either having ongoing insurance or are not satisfied with the scheme.
- Further, it is also observed that out of 174470 people, 109257 people are having vehicle age less than 1 years which means their vehicle insurance is ongoing (assuming they purchase insurance on same day as vehicle purchase)

Factors Affecting Annual Premium	
City_Code_28	0.293050083
Sales_Channel_156	-0.188717341
City_Code_48	-0.180536497
Sales_Channel_157	-0.129034336
City_Code_31	-0.116414936
City_Code_20	-0.115301895
City_Code_8	0.107127087
Sales_Channel_26	0.092305757
City_Code_1	-0.083328669
Sales_Channel_160	-0.071350149



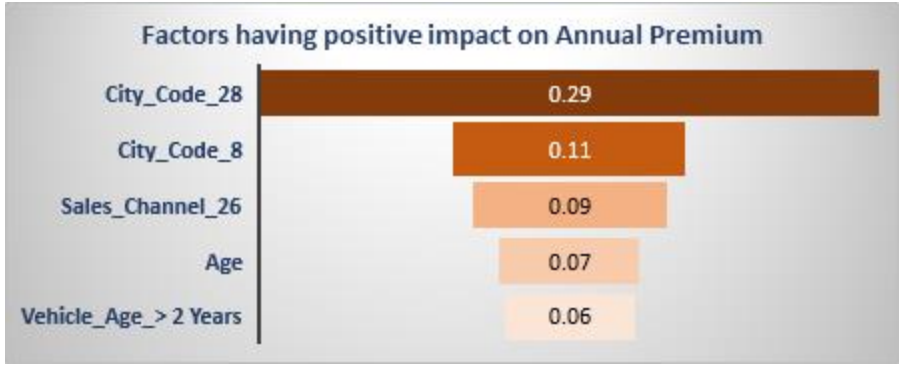
Factors affecting Annual Premium

To calculate the major factors (Top 10) affecting on annual premium, first we converted the categorical variable such as “Gender”, "Vehicle_Age", "Vehicle_Damage", "City_Code", "Sales_Channel" into binary variables leading to 216 columns.

The correlation between each independent variable and Annual premium is calculated. The top 10 affecting factors on annual premium is shown in the table.

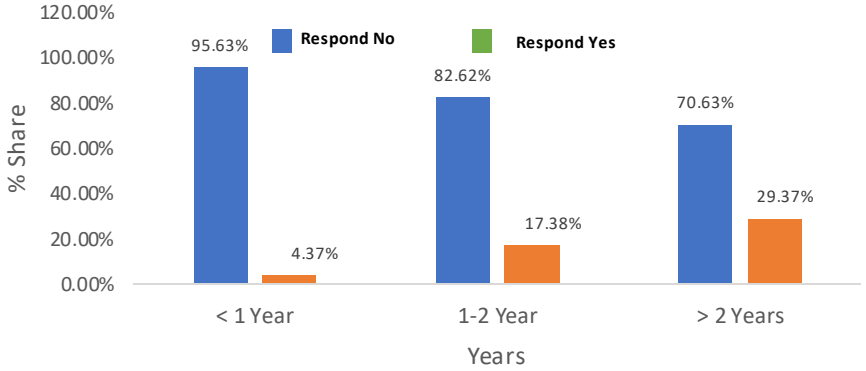
The positive correlation indicates the increase in annual premium while negative correlation indicates inverse relationship with actual premium.

Hence, to increase the revenue, the factors having positive impact on Annual Premium can be targeted to increase the sales. Hence, factors having positive impact on the Annual premium are as follows:



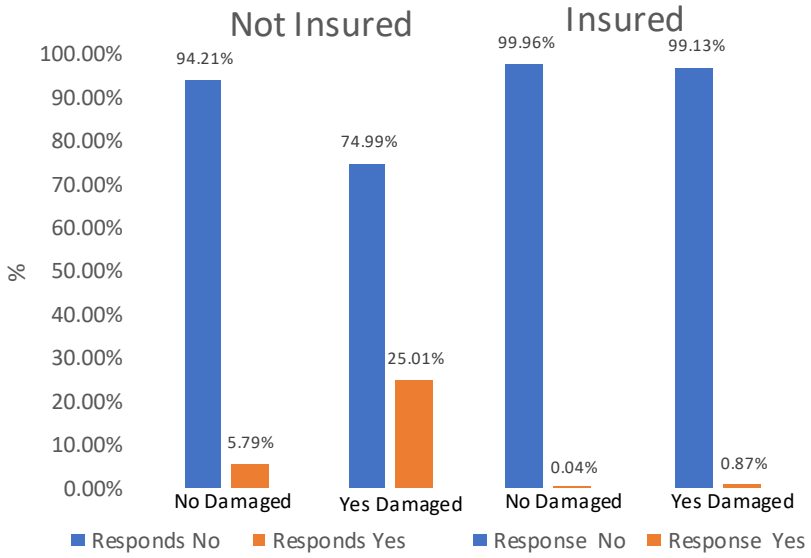
COR.html

Customer's Reponses Based On Vehicles Age



Insights: Respondents having vehicles aged > 2yrs are more inclined to buy a vehicle insurance are compared to people with vehicles aged <1 year. This might be due to in-house purchase of the insurance when the car is purchased.

Customer's Response Based On Damage Of Vehicles



Not Insured Customers' Premium vs Response		
Response	**YES**	**NO**
Mean	31620	30170
Median	33012	32065
Std Dev	18654	18194

Insights:

Customer response to Vehicle insurance (Insurance = 0) does not depend on premium to be paid as the mean ,median and Std Dev are nearly same.

Insights:

- It is evident from the above graph that the % of people who were interested in buying the vehicle insurance were higher in the case of damaged vehicle which are not insured similar trend can be seen in Insured Vehicles even though the yes response is small

Methodology Flow Chart for Predicting Response

Subsisting Data who have Insured = 0

Continuous Variable : Analysing Distribution of Each Variable and normalising if required

Categorical Variable : Creating Dummy Variable for categorical Variable

10% data for Testing and 90% data for Training

Used Smote to Make Unbalanced Train dataset Balanced

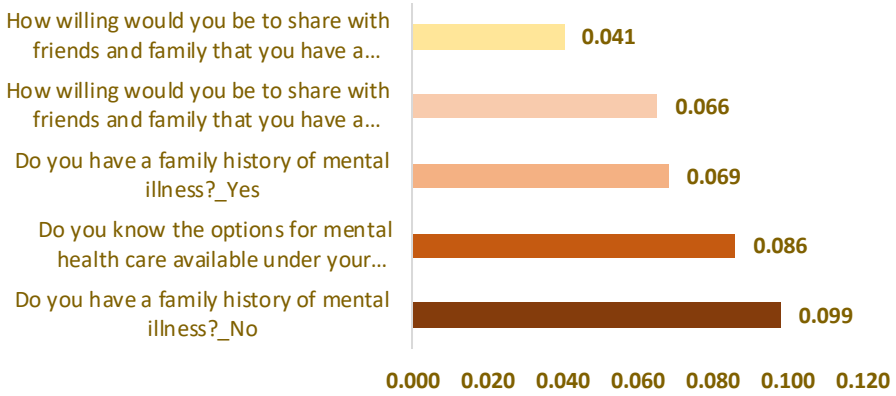
Modelling and Testing on 10% Testing Data

DecisionTreeClassifier

Specificity	Accuracy
0.851	0.708

Model → Insurance.html

Top 5 affecting factors affecting mental health illness



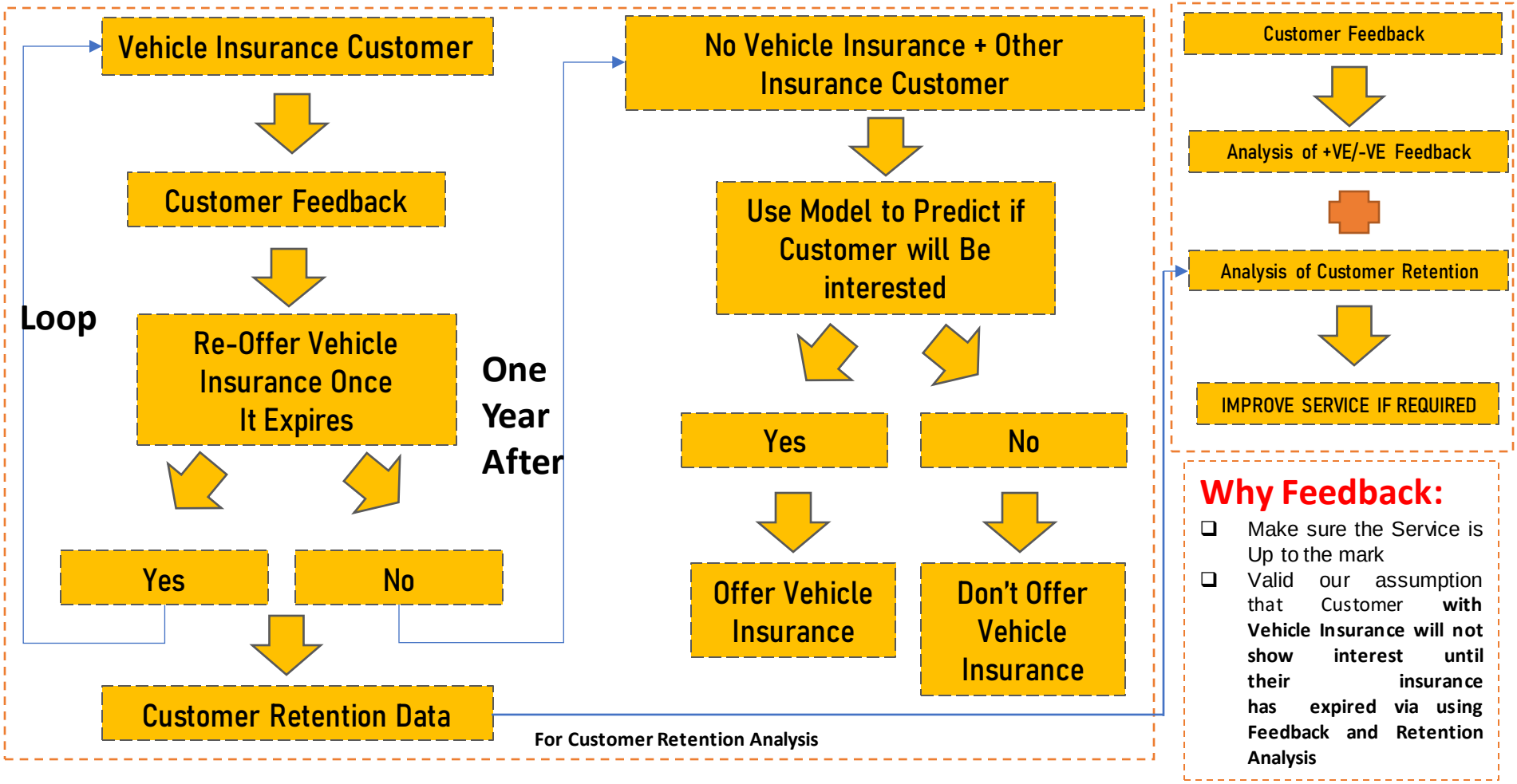
Inference

- ❑ The Model is able to predict 85.4% interested customer i.e. Response = 1 Correctly
- ❑ The Model is able to Predict 70.7% Response Correctly i.e Response = 1 or 0

Why focus on Customer with no Vehicle Insurance

- ❑ People with vehicle Insurance have shown no response (Response = 0) We are Assuming that Customer **with Vehicle Insurance will not show interest until there insurance has expired**

Strategy for Company to Apply



Why Feedback:

- Make sure the Service is Up to the mark
- Valid our assumption that Customer **with Vehicle Insurance will not show interest until their insurance has expired via using Feedback and Retention Analysis**

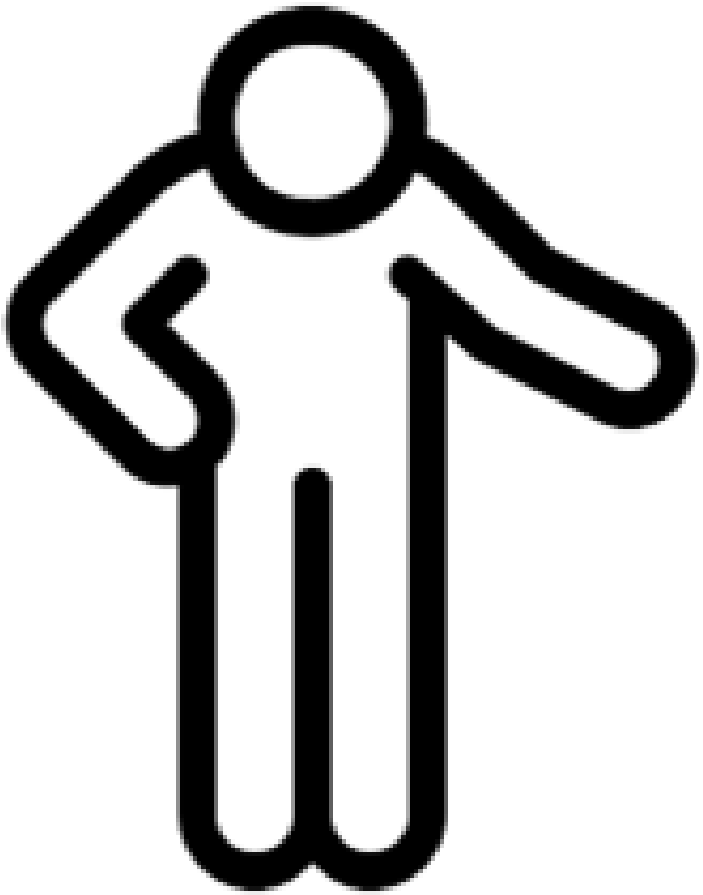
Cost Benefit Analysis of the Strategy

Benefit

- Model can help in **Identifying 85.4%** interested customer Correctly help in **increasing the Revenue** by adding **new Customers for Vehicle Insurance**
- Feedback and Customer retention analysis can help **in improving service** regarding Vehicle Insurance
- The strategy will help company to add new customer by both **improving Customer Retention** and **Increasing Customer acquisition** improving the company's **stagnate growth observed during the forecasting**

Projected Benefit of the Model

- Number of Days = 289 (Customer Association Ranges From 10 to 299)
- Expected Annual Premium from not insured interested customer (**Response = 1 and Insured= 0**) = Rs 1,471,978,440
- Expected Annual Premium for a day = Rs 1471978440 * (1/289) = **Rs 5093351**
- According to model 85% can be identify Correctly = 0.85* 5093351 = Rs 4329348
- Per Day Actual Revenue = Rs18516785 (Total Premium from Insured = 1 / 289 Days)
- Expected Increase in Customer Via Model Per day = **137**
- Expected Increase in Revenue Via Model = Rs 4329348/ Rs18516785 = **23.38%**



Thank You