

Domain Oriented Case studies

1)Telcom churn case study:

Problem statement:

In the telecom industry, customers are able to choose from multiple service providers and actively switch from one operator to another. In this highly competitive market, the telecommunications industry experiences an average of 15-25% annual churn rate. Given the fact that it costs 5-10 times more to acquire a new customer than to retain an existing one, **customer retention** has now become even more important than customer acquisition.

For many incumbent operators, *retaining high profitable customers is the number one business goal.*

To reduce customer churn, telecom companies need to **predict which customers are at high risk of churn.**

Steps performed for telecom churn case study:

- 1) Check data for null values and clean it
- 2) Check the shape and summary of data
- 3) check the 70 percentile of amount of June and July
- 4) keep only customers who have recharge 70 percentile amount and drop rest rows
- 5) Derived churn and not churn cases
- 6) Analyse the data using graph
- 7) Split the data into train and test data
- 8) Check for multicollinearity using VIF
- 9) Checked P –values for best column needed
- 10) Checked the correlation between the numerical column using heat map
- 11) model creation using train data
- 12) Model evaluation
- 13) Checked the accuracy score of model

Conclusion:

#1) Telecom company needs to pay attention to the roaming rates. They need to provide good offers to the customers who are using services from a roaming zone.

#2) The company needs to focus on the STD and ISD rates. Perhaps, the rates are too high. Provide them with some kind of STD and ISD packages.

#3) To look into both of the issues stated above, it is desired that the telecom company collects

2) E –commerce and B2B case study:

Problem statement:

Schuster is a multinational retail company dealing in sports goods and accessories. Schuster conducts significant business with hundreds of its vendors, with whom it has credit arrangements. Unfortunately, not all vendors respect credit terms and some of them tend to make payments late. Schuster levies heavy late payment fees, although this procedure is not beneficial to either party in a long-term business relationship. The company has some employees who keep chasing vendors to get the payment on time; this procedure nevertheless also results in non-value-added activities, loss of time and financial impact. Schuster would thus try to understand its customers' payment behaviour and predict the likelihood of late payments against open invoices.

Steps performed for telecom churn case study:

- 1) Check data for null values and clean it
- 2) Check the shape and summary of data
- 3) Checked for null values and replace them with 0
- 5) Derived default customers by checking whether the payment done by customer i.e receipt_date is greater than due date or not (1,0)
- 6) Merge data using customer number (here rename the column customer account number in invoice_data as customer_number)
- 6) Analyse the data using graph
- 7) Split the data into train and test data
- 8) Check for multicollinearity using VIF
- 9) Checked P-values for best column needed
- 10) Checked the correlation between the numerical column using heat map
- 11) model creation using train data
- 12) Model evaluation
- 13) Checked the accuracy score of model

Conclusion:

1) Invoice type whether it is goods or non goods has an impact on delay of payment so business team need to Consider this

2) Age of the customer can also affect the payment delay normally the people who are in the age of working Or doing in career will be able to pay the payment but rest of them can have difficulty to pay

3) BFSI Case study

Problem Statement:

As a business analyst for Home Credit, you are supposed to first gather the information and clean it to make it usable.

The bureau information is at trade level, each individual trade level information is provided. You need to apply 'Feature Engineering' techniques to roll up the information at applicant level, and thereby create manual features for model building.

Build a classification model to differentiate applicants between approves and rejects.

As a business analyst, you would want to find answers to the below questions for the bank:

How to leverage trade level information for Credit Bureaus by aggregating trade level information to applicant level in order to capture their payment behaviour?

Which application or payment behaviour factors significantly influence borrower's behaviour on any new disbursed loan?

After identifying these factors, how to leverage them in the form of a model which can be used for decisioning?

Steps performed for telecom churn case study:

- 1) Check data for null values and clean it
- 2) Check the shape and summary of data
- 3) Checked for null values and replace them with 0
- 5) Derived target customers
- 6) Merge data using SK_ID_CURR (applicant data and bureau data)
- 6) Analyse the data using graph
- 7) Split the data into train and test data
- 8) Check for multicollinearity using VIF
- 9) Checked P –values for best column needed
- 10) Checked the correlation between the numerical column using heat map
- 11) model creation using train data
- 12) Model evaluation
- 13) Checked the accuracy score of model

Conclusion:

- 1) it seems that laborers are getting issues with loans resubmission so we should give some basic facilities or in some way we need to give some easy way so that laborers can also pay the loan back in a given amount of time.
- 2) Total income is an important factor to whom loan should be given or not
- 3) we should consider someone who has filled the loan amount properly but may be one time or lesser time they make but we need to target those customers who have long number of credit due days, because they will make actual default in the loan which can affect the bank.