Sentiment Analysis:

Objective: The objective here was to find out the sentiment of the Customers writing Emails to Customer Care so as to prioritize the Email response

Context: The contact center of Flipkart receives around 100K Emails per week expressing issues with the pre or post-purchase scenarios. The Contact works in a First In First Out reactive mode and hence does not involve any sort of prioritization with respect to customer anxiety. This leads to a non-optimized way of handling customer issues leading to the most anxious customers vetting out their anger by escalating it to social media. To avoid escalation and to optimally solve the customer issues, it mandates to know the anxiety of the customer.

Methodology: The approach was to find out the anxiety of a customer based on the email text classification which defines the sentiment of the email.
While text classification was the first choice for this type of an unstructured data problem, we pinned down on the knowledge based technique of sentiment analysis specifically for this problem, which works on the premise of identifying the polarity of a statement whether 'positive', 'negative' or 'neutral' based on the unambiguous affect words like 'sad', 'happy' etc. These words based on their individual polarity and also based on combination of affect words define an overall text to be positive, negative or neutral.
While knowledge based technique of sentiment analysis a robust way for classifying text and identifying the inherent sentiment, there are two aspects to be dealt with while using this in a real world application like the one we have

1. The first one is to define what accuracy means for an unstructured text classification and how to measure it
2. The 2nd part of the problem is to design a framework to improve the accuracy

To solve for the accuracy definition part, the algorithm was run on a significant amount of data, which classified the words into their polarity and a metric was defined called the Sentiment Score which was basically the number of positive words subtracted by the number of negative words. A sample of emails from the above were manually checked and the same score was generated. The scores from the model and the manual process were compared to analyze the accuracy of the model in predicting the sentiment. This process was robust as we were comparing the customer sentiment as understood by a human agent. The first version of the model generated only about 50-55% accuracy which didn't seem fit.

The 2nd part of the problem of improving the accuracy first required us to identify what was the actual reason why the model was not able to match human performance. When looked deep into the working of the algorithm, it became apparently clear that the standard dictionary of unambiguous affect words provided in the statistical package in R was not robust enough for this context. To achieve the creation of a robust dictionary, the new words encountered during the manual calibration were tagged to their polarity and added to the existing dictionary. This process was iterated with multiple samples of emails so as to achieve a robust dictionary.

After the creation of the dictionary, the model was again run on a new sample of data and the accuracy of the data is again calibrated with the new dictionary and this time an accuracy of 80% was achieved which was good for implementation.
The model process was automated in real time with each of the emails received by Flipkart going through this model calibration and an output was generated which was basically the Sentiment Score i.e. number of positive words subtracted by the number of negative words which gives both the polarity as well as the severity of the email. The emails were then stack ranked based on the Sentiment Score and were routed to agents in that order for resolution.

Results/Impact: The prioritization of the customer emails based on the Sentiment of the customer caused the customer escalations to come down by 30% and the Customer Satisfaction Score to improve by about 300 bps

Payment Card Network Analysis:

Objective: To analyze the performance of a Payment Card Network with respect to the Success Rate of Customer Transactions

Context: An e-commerce company like Flipkart has a variety of payment options on its platform catering to the different preferences of the customer. Not all the transactions initiated by the customer become successful because of various technical issues at the Payment Gateway, Bank and Card network end. Some of these issues might be customer driven where the customer doesn't show intent to buy the product after initiating the transaction or is unaware of further procedure like OTP or PIN details but on the other hand there are non-customer dependent factors which have a potential to improve upon.

Methodology: The drivers of failure at a card network level were initially categorized into Controllable and Uncontrollable factors making it clear and robust as to the actionability The Uncontrollable factors included Customer profile dimensions like Demographics, Purchase behaviour etc. The Controllable factors featured ones which are under our actionability like routing of transactions to payment gateways and improving the efficiency of current ones
A deep Multivariate analysis was carried out across all the above drivers to pinpoint the ones which affect the Success Rate

Insights/Actions: The routing to Payment gateways was identified as the key driver. Instead of stopping at this point also strategized the potential routing scenarios for a better optimized routing

Payment Fraud Detection: