Churn Problem Statement.

We are looking at using past data to run data science analysis to predict the propensity of customers to churn soon. This would help to determine the right engagement or intervention strategy to keep customers identified as likely to churn from churning.

Simply put, the bank wants to identify the customers who are likely to have churned by the end of the incoming 180 days or six months. A churned customer is defined as one who does not have up to one customer-induced transaction on any of their accounts in a contiguous six-month period.

The bank's Data management and analytics team therefore, spooled customers' data: their demographics, accounts and transaction history over the past two years. The goal was to determine the propensity to churn for each customer.

The Analytics team spooled over eleven million dataset of the bank's customers, together with their various relationships along all the products lines. We have 2,000,000 customers' data in training set and 9,000,000 in the test set.

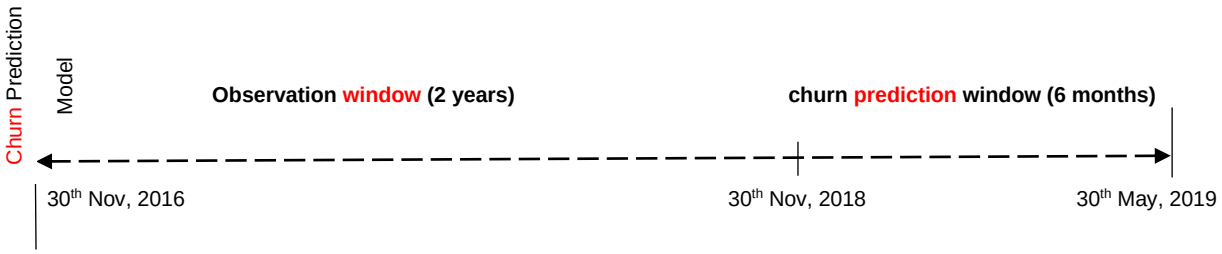
• ***We are expected to estimate the probability for each customer in test set to churn.***

Project Objectives

Project Objective	For Data Management and Analytics, the objective of the churn project is to give the bank insight through its past customer data to be able to make proactive decisions in order to retain customers and increase profitability.
Business Objective	For Access Bank, the objectives are to gain insights from its past data, and to identify customers at any stage of their lifecycle who are currently active but are likely to become inactive. This will help the bank to create a churn management strategy to assess the potential impact and forestall the customer decay process

Model Objective	The model would rank each customer between 0 and 1 on the basis of their probability to churn based on the 2 years historical data of customer behavior.
Target Base	A random 11,000,000 individual customer's data. 2,000,000 as training set and 9,000,000 as test set.

Definitions: Churn Prediction Model



- For the Churn model project, churn by definition means a consistent 6 months period of no customer initiated deposit or withdrawal activity.
- Churn here is defined as customers who did a bank transaction (either credit or debit) at least once in the Observation window and none within the Prediction window. It could also mean a cohort of customers who did not generate any revenue for the

business within a given timeframe. This definition puts a boundary to the potential churn customers that we are interested in.

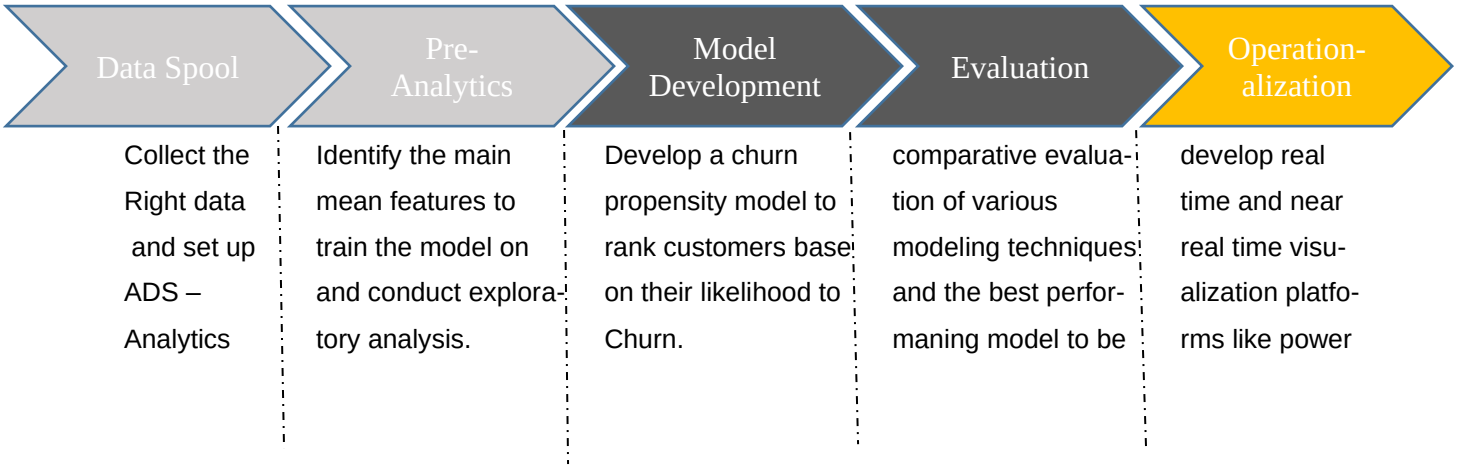
- Observation window – The customers' behavior in this window of 2 years was studied to model for the event in the Prediction window i.e. Using customer history in the last two years to determine whether they will churn in the next six months of the year.
- Prediction window – This is the period in which the event (churn or no activity) is predicted.
- The training set are the customers dataset labeled as "churn" or "not churn" while the test set are not labeled.
- Typically, from the timeline, it is easier to tell those who will churn tomorrow. But it becomes increasingly difficult as we move from left to right.

Business Analytics Approach

The methodology for model development was undertaken in 5 distinct steps.

Unlike hackathons where data scientists are required to develop a machine learning solution to a problem and upload the result to a solution checker like kaggle, Business Analytics require more. The business clients often want to understand the data from historical perspectives. And after the future has been predicted, the predictions must be described adequately for the organization to know what to do next. All of these activities will fall within the purview of the recommended Customer Value management (CVM) function.

Below is the **Data Management and Analytics Team's Approach to Business Analytics**



Datastore to selected. BI, Chat Bot.