Module 1
Retail Analytics

(Churn Analytics)

Lecture 2

16/07/2021

Customer Churn Analytics

- Churn is a term used to indicate a customer leaving the service of one company in favor of another company.
- Customer churn analytics is the process of tracking attrition in a given length of time.
- It helps us to predict customer churn in advance and taking corrective measure or indirect action before we lose the customers to competitors.

Classification of churn models

- In a real world application field, it is not the easy way to find a unique model which is appropriate for designing churn solution.
- A predictive churn model gives valuation about churn probability in forthcoming period of time. Calculate probability that certain customers will commit churn in forthcoming period of time.
- Time dependent churn models includes time component which is critical for the final model construction.

- Segmentation churn models attempts to find related attributes, with assumption that the fact that within churners exists differentiation. Also if the sample segment falls within churner population, it is possible to create different predictive churn models for each segment.
- Profiling churn models : Profiling models gives complete picture of typical churner as member of churner population or as a member of the segment. Profiling analysis include variety of data mining techniques.

- Churn classification model

❖Soft churn models are associated with portfolio where the user started to reduce the product usage or their service.

❖Hard churn models is if the client/buyer/user broke the contract. In this case variety models can be applied for finding the causes of churn, or patterns, or segments and profiles as well.

- Churn prevention models help in situation where churn is not still identifiable or problem in company. These models are often established without particular information about future market conditions and it could be based on refined guesses and expected market movement.

Churn Solution Modelling areas and data mining techniques

- Models for churn migration are models developed and applied mostly where churn become obvious and is a big problem within company.
- The models depend on predictive churn models, segmentation models and time dependent models.

Predictive Modelling

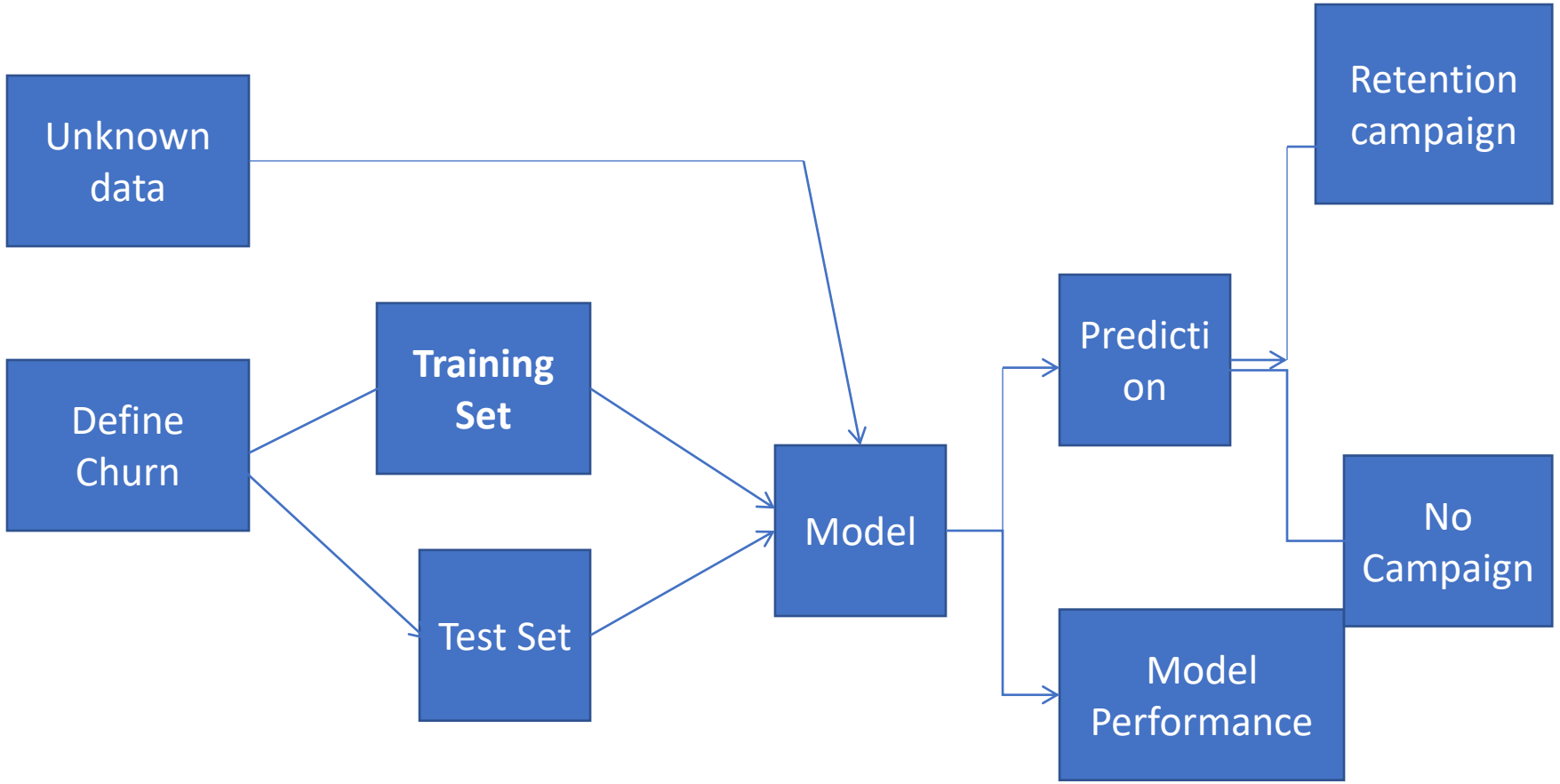
- Linear Regression
- Logistic Regression
- Neural Networks
- Bayesian Networks
- Decision Trees

Segmentation

- K means Clustering
- Support Vector Machines
- Self Organizing Maps

  Time Dependent Models
- Cox Regression
- Kaplan Maier Models

Churn Prediction Process



Churn data modeling

Churn data

- The dataset contains 20 predictors worth of information about 3333 customers along with the target variable churn ,an indication of whether that customer churned (left the company).
- To begin, it is often best to simply take a look at the field values for some of the records.
- The response variable *churn* is a flag variable having two values, *True* and *False.*
- Summary, Histograms or bar charts etc.
- Develop a model of the type of customer likely to churn.ie, we are interested in predicting whether a customer would leave the cell phone company's service (churn) based on a set of predictor variables.

The variables in the dataset

- State: categorical for 50 states and the district of Columbia
- Account Length: Integer valued, how long account has been active
- Area code : categorical
- Phone number: essentially a surrogate for customer ID
- International plan: Dichotomous categorical,(yes/no)
- Voice mail plan : Dichotomous categorical,(yes/no)
- Number of voicemail messages: Integer valued
- Total day minutes: Continuous, minutes customer used service during the day
- Total day calls: Integer valued
- Total day charge: Continuous, based on above two variables

- Total Eve Minutes :Continuous, minutes customer used service during the evening
- Total Eve calls: integer valued
- Total Eve charge :Continuous, based on the above two variables.
- Total Night Minutes :Continuous, minutes customer used service during the night
- Total Night calls: integer valued
- Total Night charge :Continuous, based on the above two variables.
- Total International Minutes :Continuous, minutes customer used service to make international calls
- Total International calls: integer valued
- Total International charge :Continuous, based on the above two variables.
- Number of calls to customer service: Integer valued
- Churn : Target, Indicator of whether the customer has left the company.

To do....

- Missing data imputation and data cleaning
- Use a graph to visually determine whether there are any outliers among the number of calls to customer service?
- Identify the range of customer service calls that should be considered outliers using
- (a) Z-score and (b)IQR method
- Transform the day minutes attribute using Z-score standardization
- Work with skewness as follows
- (a) Calculate the skewness of day minutes
- (b) Then calculate the skewness of the Z-score standardized day minutes. Comment
- Construct a normal probability plot of international minutes