CUSTOMER REPURCHASE ANALYSIS IN B2B CONTEXT: A BAYESIAN HIERARCHICAL FRAMEWORK

Author : Anindya Sankar Dey, Dr. Jayanta Kr. Pal, Subhasish Misra, Abraham Paul

HP GBS Analytics

Speaker : Jerry Shan, HP Labs

MARCH,2011

AGENDA

- INTRODUCTION
- METHODOLOGY
- KEY RESULTS
- ADVANTAGE & LIMITATIONS

INTRODUCTION

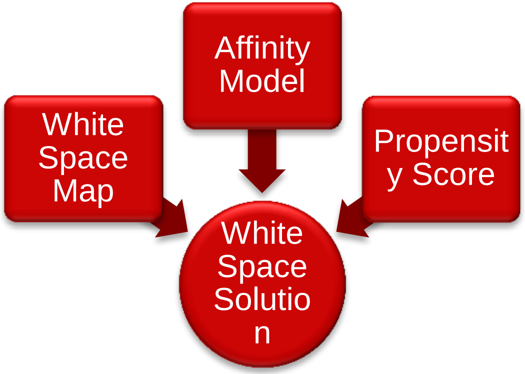
BACKGROUND

- HPSS Sales Team want to have a complete White Space solution.
- The analytics team is already building a tool to address the affinity model part of the solution
- It is easier to target existing customers base

Objective

- To develop an algorithm that will help to assign a propensity score i.e. probability to transact in the future to each HPSS account in the different sales play.
- Higher score implies that an account is more inclined to buy HPSS product in a specified time period.

BUSINESS BENEFIT



- Business can target those customers who are more likely to transact with HP in the future
- The Affinity model along with the propensity score will help in targeting most probable customers at a PL level
- The propensity scores will be a valuable add-on to the TARGET tool and will make it a White Space Solution Tool.

METHODOLOGY

METHODOLOGY

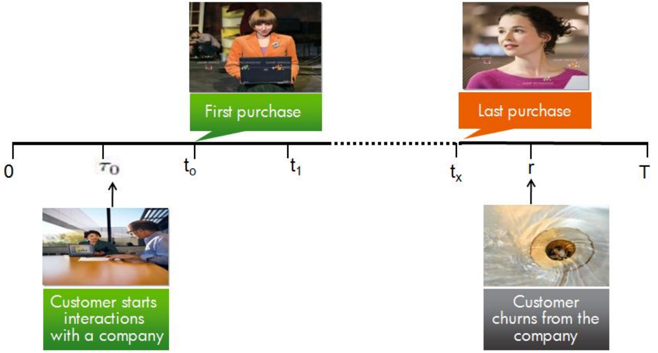
- To compute the propensity score we need two parameters:-
  - The probability to churn from the company – ‘p’
  - Average Transaction per month by a customer – ‘λ’

- Transaction at each time point is considered to follow a Poisson Distribution with mean λ

- Compute the parameter estimates as follows
  - Consider Prior Distribution for each parameter
  - Compute their Posterior distribution
  - Use Markov Chain Monte Carlo – Gibbs Sampler simulation technique.

- Compute the propensity score for the next k time points given by $(1 – p)*(1 – e^{–k\lambda})$ using the obtained parameter estimates.

MODEL FRAMEWORK



ASSUMPTIONS

- The customer starts interacting with a company some time before its first transaction in the dataset.
- There is a time gap between a customer’s last recorded transaction and when the customer churns.
- Above assumptions give rise to the two latent variables r and τ_0 which are used to estimate the parameters of the model.

PRIOR AND POSTERIOR DISTRIBUTION

Definition

In Bayesian statistical inference, a **prior probability distribution**, often called simply the **prior**, of an uncertain quantity p is the probability distribution that would express one's uncertainty about p before the "data" is taken into account. It is meant to attribute uncertainty rather than randomness to the uncertain quantity. The unknown quantity may be a parameter or latent variable.

One applies Bayes theorem, multiplying the prior by the likelihood function and then normalizing, to get the **posterior probability distribution**, which is the conditional distribution of the uncertain quantity given the data.

Parameters of prior distributions are called **hyperparameters**, to distinguish them from parameters of the model of the underlying data.

Prior Distribution Assumed

- $\lambda \sim$ Gamma(k,θ)
- p ~ Beta(a,b)
- r ~ Uniform(0,T)
- $\tau_0 \sim$ Uniform(0,T)

MCMC AND GIBBS SAMPLING

Markov Chain Monte Carlo

Markov chain Monte Carlo (MCMC) methods are a class of algorithms for sampling from probability distributions based on constructing a Markov chain that has the desired distribution as its equilibrium distribution. The state of the chain after a large number of steps is then used as a sample from the desired distribution.

Gibbs Sampling

Definition: Gibbs Sampling is a specific type of MCMC method.

When we use it: Suppose we have a bivariate random variable (x , y) and we wish to compute one or both marginals p(x) and p(y). Idea behind the sampler is that it is far easier to consider a sequence of conditional distribution, p(x | y) and p(y | x), than to obtain marginals by integrating the joint density p(x , y) ,e.g. $p(x) = \int p(x , y)dy$.

Description: The sampler starts with some initial value of y_0 for y and obtains x_0 by generating a random value from the conditional distribution $p(x \mid y=y_0)$. The sampler then uses x_0 to generate a new value of y i.e. y1 from $p(y \mid x = x_0)$. The Sampler then proceeds as follows:-

$$xi \sim p(x|y=y_{i-1})$$

$$y_i \sim p(y|x=x_i)$$

Repeating the process k times generates a Gibbs Sequence of length k. Initial part say first m points of the sequence is left out as burn in periods (period after which the sampler becomes stable) and from the rest a subset of points(x_j , y_j) for ≤j≤l<k-m are taken as our simulated draws.

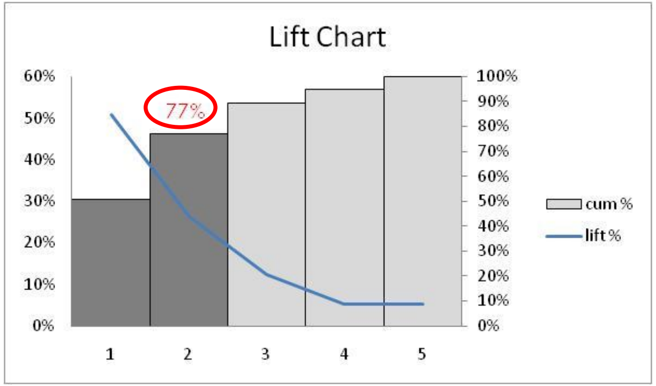
KEY RESULTS

Classification Table (Predicted =1 if score >= 0.65 Predicted =0 if score < 0.65)		Predicted			
		0		1	
Observed	0	333	58%	182	32%
	1	16	3%	41	7%

– ANALYSIS: In APJ analysis was performed for existing accounts of HPSS business for affinity

– RESULT:

- 72% of actual repurchases in out of time validation sample was predicted correctly
- In validation dataset top 40% customer captured 77% of actual repurchases.
- Model has a concordance of 76%
- Rank Correlation between probability of customers to purchase in next 1 year and next 2 year is 0.82.



IN CONCLUSION

ADVANTAGES	LIMITATIONS
– Only transaction history of a customer is required to build the model . – Random fluctuation in data can be easily eliminated from modeling as no dependence on various variables. – Instead of bucketing, we compute individual propensity/scores. – Forward looking solution: Relative ranking of customers based on their propensity scores might change with time. – **Model can be replicated for other Sales play, PL , BU, etc. It is a solution for problems of similar nature**	– Complex model. Need to understand statistical techniques for implementing the model. – Some times external factor influences repeat purchase behavior of customers which is not captured by this model though it was not significant in our scenario. – The entire modeling exercise is difficult to automate as manual inputs and intervention is needed in some steps. – Need substantial computing power for large datasets.