Paper 8481-2016

Make predictions more predictive. Lift retail marketing campaign response by blending Logistic Regression and Cox Proportional Hazards Model

Hsin-Yi Wang MS., Alliance Data Systems, OH

ABSTRACT

Logic model produced propensity scores have been intensively employed to assist direct marketing name selections. By which, only customers with absolute higher likelihood to respond are mailed offer to in order to achieve cost reduction, thus event ROI is increased. A fly in the omelet is, compared to model building performance time window, usually 6 month to 12 month, a marketing event time period is usually much shorter. As such, this approach lacks of the ability to deselect those who with high propensity score but unlikely to respond to an upcoming campaign. Considering to dynamically building a complete propensity model for every upcoming camping is nearly impossible, incorporating "time to respond" has been of great interests to marketer to add dimension to enhance response prediction. Hence, this paper presents an inventive modeling technique combining Logistic Regression and Cox Proportional Hazards Model. The objective of the fusion approach is to allow a customer's shorter "next to repurchase time" to compensate for his/her insignificant lower propensity score in winning selection opportunities. The method is accomplished using Proc Logistic, Proc Lifetest and Proc Phreg on the fusion model building in SAS environment. This paper also touches how to use the results to predict repurchase response by demonstrating a case of repurchase prediction on 12m inactive customers of a big box store retailer. At the end, this paper will share results c omparison between fusion approach and logit along. Comprehensive SAS® codes will be provided in the appendix.

INTRODUCTION

In the US, it is not surprising that major retailers leverage predictive models to drive marketing targeting solutions to decrease mailing size and increase marketing response rate. The most event-targeting fundamental prediction objective is how likely this cu stoemr is coming to shop in the future. The approach is to encourage customers' natural intention, which means, the more the propensity to shop of a custoemr is, the more target value this customer has. Thus, marketing modelers develop and improve propensity to shop models to score customers and based on the scores, to provide recommendation list of customers to send marketing communication pieces to depends on various event type and what kind of customer behavior this event is specifictedly trying to derive. The propensity to shop model methodology is matured and well applied adopted logistic regression. In most cases, instead of "Mail-to-All", only customers have scores higher than average will be considered worthiness for retailer spend extra to market with confidence while those with low scores are being weeded out as they are expected to not respond and offers mailed to them may not drive postiitve return of investment. Therefore, propensity to shop model saves cost on ineffective targeting to increase response rate, profit and ROI especially for retailers with economic scale or heavily promotional culture. This paper is aimed to discuss how to tune the approach of "propensity to shop" prediction and lift response to drive better incremental sales for business.

Logit Model and the Limitation

Logistic regression has reputation for its powerful dichotomous response seperation capability, easiness to for business communications and performance robustness. Most of marketing predictive models either leverages it fully or at least significant partially. Thanks to data technology advancement and innovation for the past decade, comprehensive data such as demographic, lifestyle, geographic and web activities has been well collected for marketing quantitative research. However, the lift of prediction has still yet satisfies marketers and marketing modelers. According to DMA, direct marketing association report: in average marketing event targeting predicted performance, for one camapgin, the best group of customers can achieve 5% to 6% response rate, and 3.7% response rate with house list while random selection may only yield 1% (DMA research report, 2015). Compared the statistics a decade ago, transactional information was the only resource and the average response rate was in average 2% with random selection is below 1.5 %(DMA report, 2012). Given the challending fact, marketers have not stopped studying : How can we bring up response lift to make our prediction more predictive.

To ensure perfroamcne stability, often propensity to shop take equal length of observation and prediction time window, for retail industry, usually 12 month is the length being chosen. Which says that, 12 month of customer historical data is used to predict the same customer's future 12 month purchasing behavior. Given this fact, we know a marketing campaign is most of the time has execution time window far less than prediciotn time window, for example, two weeks before a holiday or even just a weekend. While it can be mathematically proven 12 month propntiy equates any units time length propensity subject to all unit time propensity to shop is mutually independent, however we know this is usually not the case. thus, there is a wigger room of some customers may have higher propensity to shop in the next 12 month but not nessasary have higher propensity to shop this weekend.

Can massive model building process lifts the prediction performance? For example, build a model for evenry campaign or every month? It is true that some marketers do this, but other problems arise: the ignorance of seasonality, the fact of covering less pruchaisng senerio and the inefficient allocation of moldeler repsuces. In fact, the yield recommendation list does not shown significant lift with massive modeling, moreover, retailers with smaller customer base cannot afford such intensive model buildings human resources and this also makes model performance track more difficult using different models. An opportunity to shed light on the prediction is to incorporate the "time to event" prediction. With not only propensity being considered, we also take a deep look into individual's immediate response probability or time to next purchase. The expectation for adding this dimension is to compensate the 12 month view restriction of logic model.

Survival Modeling

Survival analysis is also known as time to event analysis with focus of studying the relationshop of event occurrence and elapsed time. Survival modeling was developed and applied in clinical field for patient lifetime estimation therefore survival rate over time. Later, its predictive capability of modeling time to event has caught interests to markers. Until now, many marketing studies have been done leveraging survival analysis as well as by various domains of industries. Most marketing studies leveraging survival analysis weight more focus on identifying causing factors which impact event occurrence rate. For example, customer churn, attrition, and tenure have had paper discussed as the event in survival study. However, this paper will address less in this direction. Instead, individual immediate response liekihood predition which is part of survial modeling derivatives is the purposes the discussion in this paper. In following, this paper is going to illustrate that, by blending logistic model and survival model, how marketers can optimize prediction accuracy to improve marketing event name selection as well as response rate.

- Hazards

Hazards in survival modeling is served as model response for estimate survival propability. Up until an event occruance, the higher the hazards, the more likely the event is going to take palce. For example, regular people's hazards factor to lung cancer can be gender, smoking behavior indicator and residential air population measurement. Below is a brief of introduction of survival model essential formula and how it derives the survival probability.

- **The Survival Function**

$$S(t) = S(t-1) * \left(1 - h(t-1)\right) \; ; \; S(0) = 100\%$$

- **The Hazard Function**

$$\lambda(t) = \lim_{\delta \to 0+} \frac{P(t \le T < t+\delta \mid T \ge t)}{\delta} = \frac{P(t \le T < t+\delta)}{P(T \ge t) \times \delta}$$

$$= \lim_{\delta \to 0+} \frac{S(t) - S(t+\delta)}{\delta} \times \frac{1}{S(t)} = \frac{f(t)}{S(t)}$$

$$h(t) = \frac{f(t)}{F(t)} = \frac{Number\ of\ Customers\ Left\ in\ Time\ T}{Number\ of\ Total\ Customers\ in\ Time\ T}$$

$\lambda(t)$ is the failure rate being modeled as the event's probability to occur at time T which is larger than time t and less than t plus a very small time interval of delta. Here, survival implies an entry's probability to survive until time T. All entries were alive when survival experiments or studies begin.

- **Cox's Proportional Hazards Regression Model**

$$\boldsymbol{h(t|X) = h_0(t)e^{(\beta_1 X_1 + \beta_2 X_2 + \cdots \ldots + \beta_n X_n)}}$$

Proposed by Sir David Cox, Cox's Proportional Hazards Regression Model is a semi-parametric model which allows no density function of parametric assumptions of time (t). It also assumes hazards are proportional over time. This technique is chosen because the study focus is to predict immediate response likelihood for retail customers given historical purchasing behavior. This can be interpreted as hazard rate in statistically term.

RETAIL CUSTOMER LIFECYCLE AND SURVIVAL

This paper will demonstrate the modeling process with a real business case. For illustration purposes, relevant business terminology is defined and declared in this section.

In retail business, it is common customers are being segmented by their brand shopping life stage. Generally, within first few month of initial purchase, the customers are defined as "New". And if a customer can continue in business of by maintaining their purchase gaps less than X month, they are considered "Live" customers which is the most profitable and stable customer lifecycle segment. As soon as a customer passes X month since last purchase, they are being defined as "Dead" where X is agreed by industry a typical customer lifetime. The term "Dead" may not necessarily mean those customers will never shop back but really to convey that these customers are much more difficult to "wake up" for repurchase than those "Live" ones. 12 month is under consensus for most retailers as X. Please note, X differs by customers and by business kind as well. For example, some customers may be in danger of being churn once time since last purchase gets close to 6 month. And there are also business has longer customer lifecycle such as furniture stores. The model is going to be demonstrated in the paper can apply to most business kind as long as the X is chosen appropriately.

The demonstration of this paper is done through propensity to shop case study of Retailer M. M is a national known big box retailer which sells a variety of products from kitchenware to shoes and from clothes to cosmetics. A logistic regression method based model is available for M in the use of monthly marketing campaign targeting. As the retailer are sending out millions of direct campaign each quarter. With this economic scale, significant response lift can mean significant profit increase or cost reduction.

<u>DLP_to_current</u>- The term describes of "Days Since Last Purchased". This term describes the gap between observation end time and last purchase date. In general, the longer the gap, the more likely of this customer leaves the brand increases. This value is calculated in month.

<u>LAPSED</u>- a equivalent term to "Dead" in shopping cycle. In our case, if a customer of M has DLP_TO_CURRENT longer than 12 month, we label this customer as being "lapsed". In retail, an average lapsed customer has significant lower marketing pieces response rate and higher cost to acquire for campaigns compare to active ones.

<u>OBSERVED</u>- If a customer's Active status can be explicitly defined they are being labeled as observed case.

SURVIVAL MODEING PROCESS FOR TIME TO NEXT PURCHASE PREDITCITON

Study Objective

To demonstrate how to build customer to time t survival prediction model using Cox Proportional Hazards Model and evaluate the lift performance blending survival model with propensity model.

Methodology

Traditionally, the entry lifetime starts from experiment begin. As in marketing world, longer tenure with business actually indicates higher loyalty and least implies churn increase which contradicts the design of survival. The innovation is to flip the lifetime by locating an observation anchor time and track back throughout lifetime to capture the trend of survival pattern. In the modeling set up, we will do another transformation for prediction. This will be further discussed in modeling section of this paper.

Data Sources

Retailer M is a national well know big box store. They have millions of active customers and for those M collected individual level Demographic data and the Transactional data. These are being used for this study. In our modeling, more than 50 variables are directly calculated or indirectly derived from the above two database. Some sample variables are provided in the following:

Demographic: Age, Income, Education, and Gender

Transactional: Spend in dollar, units of item purchased, number of transactions.

Sampling Strategy

Among the millions of customers who shopped with retailer M during the year of 2015 are being captured as our modeling base. Their activities are then followed back by at most 12 month to identify the gap to previous purchases as well as to confirm “Live” status. There are about 10% among all customers who are unable to confirm “Live” status due to lack of initial transaction date which makes their status potentially fluctuate between “New” or reactivated from “Dead”. In survival analysis, these customers are censored observations. Their behavior will be considered in Live behavior but their leave of study will not discount the survival rate.

Pre-Modeling Preparation

- **Explanatory Data Analysis**

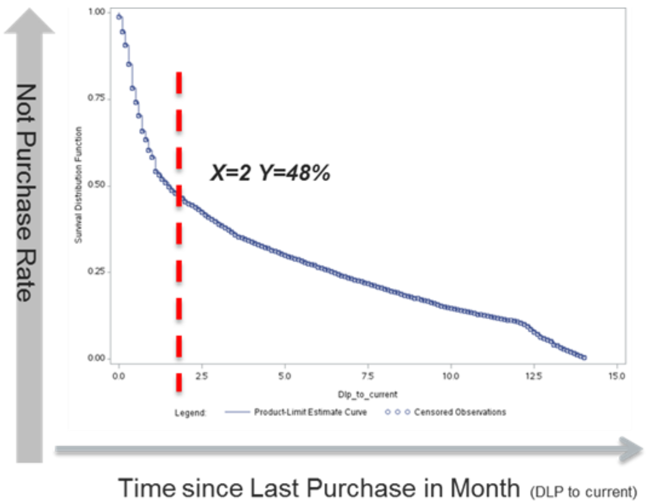
There are preparative steps taken to purify data and control outliers to pertain model performance stability. Those include and not limited to check missing rate for each of the variables, to discard variables with higher than 30% missing rate, to run univariate exploration and to perform distribution examination against each of the variables. Those steps are performed for this study. Furthermore, if 500% or more indifference is captured between variable Max and the 99 percentile, the variables are then to be capped with 99.5 percentile values. Appropriate imputations are also applied to replace missing. If missing rate is less than 5%, the values are replaced by the median. If missing rate is greater than 5%, a distribution raised from valid existing values is calculated and applied to replace missing. There are no categorical variables in this study. Only binary and continuous numeric variables are created for this study. Using Transaction data, we can calculate DLP_TO_CURRENT in month with decimals. For those who has DLP_TO_CURRENT as missing, 0 is coded for their values.

We use ***PROC lifetest*** to examine the survival pattern by DLP_TO_CURRENT by taking a sample to construct the Kaplan-Meier Curves plot. By doing this, we can re-examine assumption of the survival relationship constructed by not-purchase-rate against each day since last purchase time point.

In the following is the sample code used to plot the Kaplan–Meier curve using retailer M data.

```
%let status_var=observed;
PROC lifetest DATA=sample_perc plots=survival(atrisk cb) outs=suvout;
time dlp_to_current*&status_var.(0);
run;
quit;
```

The chart below illustrates the survival probability from purchase occurrence to retailer M customers.



Observation Base
14m active customers with 0 to 14 mon. since last purchase

Interpretation of Curve
Within **X** month since last purchase, the non-purchase rate is **Y**

Censoring Case
New customers who have only 1 purchase record

If we would like to further exame the survival curve by group, the strata option can prefrome the task. for exmaple, we may further define customers by their spend. Customers have more than overall average spend are labeld 1 in OverAvgSpend indicator. The curves look similer as higer spenders are more sensitive to DLP. Definitely the value customers behavior change is worth more attention.

Variable Selection

Final short list of variables candidates is stored in macro variable "&suv_var_temp." Proc PHREG is adopted to perform stepwise model selection. The model developed with optimal selection criteria is saved by ODS output statement as well as the OUTTEST option.

Building Cox Proportional Hazard Model Process

- Modeling Codes

```
ods trace on;
ods output ParameterEstimates=est_temp;
PROC PHREG DATA =mdata_train_f OUTEST =model_temp plots=survival;
MODEL DLP_to_current*&status_var.(0) = &suv_var_temp. / SELECTION = STEPWISE SLENTRY =
0.05 SLSTAY = 0.05 DETAILS;
baseline out=base survival=surv1;
output out=s1out survival=surv ld=exp  xbeta=Xb resmart=Mart resdev=Dev  logsurv=ls loglogs=lls;
run;
ods trace off;
```

- Scoring Codes

```
PROC PHREG data=&scoring_base._f inest=mysvmodel;
MODEL DLP_to_current*&status_var.(0) =&suv_var_fin.;
output out=s1out survival=surv ld=exp  xbeta=Xb resmart=Mart resdev=Dev  logsurv=ls loglogs=lls;
run;
quit;
```

In this procedure, &status_var. macro variable equates “Observed” where 1 indicates true “Live” and 0 implies “Censored” for which we have no further knowledge to tell the status. The final model provides proportional hazard function and produces survival score to time T which is our observation anchored.

The final model has 9 variables. The factors cover trips taken “Lively”, age, discount received, numbers of items purchased and payment behavior.

Model Prediction and In-time Validation

Cox Proportional Hazards Model estimates survival probability to time T. Using the score, we can rank customers in to buckets. In our study, the retailer M customers are assigned to one of 10 buckets in order based on their survival probability. The following shows the rank deciles and average DLP_TO_CURRENT values for each rank. Rank tie does not exist and a good probability distribution spread is observed. We will do more copies of validate ion to test the stability. In the same modeling time window, we used the built model scorecard to assign score on another mutually exclusive sampled population and the performance agreed to direct model results. The top two deciles have nearly identical survival estimates. And continue downward, the survival probability drops significantly once DLP_TO_CURRENT starts to increase.

hazard rank	Survival Probability
0	8%
1	21%
2	32%
3	42%
4	51%
5	60%
6	69%
7	76%
8	84%
9	95%

BLEDING RESULTS REVIEW BY LOGISTIC MODEL AND TIME TO NEXT PURCHASE MODEL AND NAME SELECTION STRATGY ALTERNATIVE RECOMMENDAITONS

Recall that retailer M owns a predictive model which is used for momently campaign targeting. As the true redemption data is not applicable, an inferred response is calculated based on their activity for model performance assessment. If a customer is captured repurchase in the following month after data scored, we recognize the customer is as “responded”.

Next, we layer the two predictions of propensity predictive and survival probability by scoring each of more than 1million customer dataset with each model. As such, a customer will have a survival score and a propensity score. The higher the survival score means the higher change a customer to survive until event. Same logic applies, the higher the propensity scores the higher likely this customer will respond in the future 12 month time window. We create a population density table by propensity score rank and by survival score rank.

The table value is the percent of customers repurchase in the following month after scoring in percentage. Typically, 40% of customers are being targeting recommended as they will yield over 100% gain off random selection based on model. Thus, we highlighted the cells contains top 40% of customer size with light red color.

The blending of models is done by logistic and survival model overlay. Each 12 month active customers in obsevation is assigned two model scores. The immidiate response rate is simulated by bi-rank coming back to shop rate. It is calcuated by assigning cell percentage represents the counts of two model rank overlay divided by the total one month after scoring come to shop population. The following results reveals consistent Pattern appears in the most recent months after scoring.

First+Second Mon. after Scroing		Hazard Rank										
		0	1	2	3	4	5	6	7	8	9	Grand Total
Logistic Rank	0	1.77%	2.05%	2.00%	1.98%	1.74%	1.85%	1.89%	1.79%	2.10%	4.43%	21.60%
	1	1.94%	2.05%	1.98%	2.10%	1.78%	1.81%	1.80%	1.58%	1.51%	1.18%	17.74%
	2	1.86%	1.77%	1.87%	1.86%	1.51%	1.60%	1.58%	1.29%	1.03%	0.47%	14.85%
	3	1.58%	1.55%	1.56%	1.47%	1.33%	1.36%	1.26%	1.03%	0.86%	0.31%	12.32%
	4	1.31%	1.25%	1.19%	1.27%	1.07%	0.97%	1.00%	0.94%	0.75%	0.30%	10.04%
	5	1.07%	0.84%	0.89%	0.94%	0.65%	0.62%	0.74%	0.85%	0.84%	0.38%	7.82%
	6	1.17%	0.73%	0.74%	0.49%	0.28%	0.38%	0.44%	0.53%	0.63%	0.67%	6.06%
	7	0.42%	0.67%	0.53%	0.45%	0.40%	0.31%	0.20%	0.29%	0.34%	0.58%	4.19%
	8	0.77%	0.28%	0.34%	0.25%	0.40%	0.43%	0.30%	0.24%	0.17%	0.28%	3.46%
	9	0.02%	0.02%	0.02%	0.02%	0.09%	0.17%	0.27%	0.37%	0.51%	0.42%	1.92%
Grand Total		11.91%	11.22%	11.12%	10.84%	9.25%	9.49%	9.48%	8.93%	8.73%	9.02%	100.00%

(a-1)

Second Mon. after Scroing		Hazard Rank										
		0	1	2	3	4	5	6	7	8	9	Grand Total
Logistic Rank	0	2.08%	2.34%	2.23%	2.22%	1.93%	2.05%	2.07%	1.97%	2.30%	4.99%	24.19%
	1	2.07%	2.18%	2.02%	2.14%	1.82%	1.83%	1.81%	1.59%	1.50%	1.21%	18.18%
	2	1.94%	1.81%	1.86%	1.84%	1.50%	1.56%	1.53%	1.25%	1.00%	0.47%	14.76%
	3	1.64%	1.54%	1.52%	1.43%	1.25%	1.27%	1.19%	0.97%	0.81%	0.31%	11.93%
	4	1.33%	1.22%	1.17%	1.19%	1.00%	0.90%	0.92%	0.88%	0.71%	0.29%	9.61%
	5	1.11%	0.81%	0.81%	0.85%	0.60%	0.55%	0.66%	0.79%	0.78%	0.35%	7.31%
	6	1.16%	0.69%	0.68%	0.45%	0.25%	0.34%	0.40%	0.47%	0.56%	0.62%	5.64%
	7	0.38%	0.61%	0.43%	0.39%	0.36%	0.28%	0.18%	0.26%	0.29%	0.51%	3.69%
	8	0.77%	0.28%	0.31%	0.22%	0.36%	0.36%	0.26%	0.21%	0.15%	0.23%	3.16%
	9	0.01%	0.02%	0.02%	0.02%	0.09%	0.16%	0.23%	0.31%	0.40%	0.26%	1.52%
Grand Total		12.50%	11.51%	11.05%	10.75%	9.18%	9.30%	9.25%	8.71%	8.50%	9.25%	100.00%

(a-2)

		Hazard Rank										
	Row Labels	0	1	2	3	4	5	6	7	8	9	Grand Total
Logistic Rank	0	1.99%	2.19%	2.16%	2.17%	1.90%	2.05%	2.09%	1.96%	2.33%	5.08%	23.94%
	1	1.89%	2.05%	2.01%	2.11%	1.80%	1.84%	1.84%	1.63%	1.58%	1.25%	18.00%
	2	1.79%	1.72%	1.83%	1.82%	1.50%	1.60%	1.61%	1.28%	1.01%	0.47%	14.63%
	3	1.49%	1.47%	1.50%	1.42%	1.30%	1.35%	1.24%	0.99%	0.81%	0.29%	11.87%
	4	1.22%	1.19%	1.11%	1.24%	1.02%	0.94%	0.95%	0.87%	0.68%	0.28%	9.51%
	5	0.99%	0.79%	0.86%	0.90%	0.61%	0.59%	0.70%	0.79%	0.75%	0.34%	7.32%
	6	1.08%	0.68%	0.71%	0.47%	0.27%	0.36%	0.41%	0.49%	0.58%	0.56%	5.62%
	7	0.39%	0.63%	0.53%	0.42%	0.36%	0.28%	0.20%	0.28%	0.33%	0.53%	3.93%
	8	0.71%	0.26%	0.32%	0.23%	0.35%	0.40%	0.29%	0.23%	0.17%	0.27%	3.23%
	9	0.02%	0.02%	0.02%	0.02%	0.08%	0.17%	0.26%	0.35%	0.50%	0.50%	1.95%
	Grand Total	11.57%	11.00%	11.05%	10.83%	9.20%	9.58%	9.59%	8.88%	8.74%	9.57%	100.00%

(a-3)

First Mon. after Scroing		Hazard Rank										
		0	1	2	3	4	5	6	7	8	9	Grand Total
Logistic Rank	0	1.99%	2.19%	2.16%	2.17%	1.90%	2.05%	2.09%	1.96%	2.33%	5.08%	23.94%
	1	1.89%	2.05%	2.01%	2.11%	1.80%	1.84%	1.84%	1.63%	1.58%	1.25%	18.00%
	2	1.79%	1.72%	1.83%	1.82%	1.50%	1.60%	1.61%	1.28%	1.01%	0.47%	14.63%
	3	1.49%	1.47%	1.50%	1.42%	1.30%	1.35%	1.24%	0.99%	0.81%	0.29%	11.87%
	4	1.22%	1.19%	1.11%	1.24%	1.02%	0.94%	0.95%	0.87%	0.68%	0.28%	9.51%
	5	0.99%	0.79%	0.86%	0.90%	0.61%	0.59%	0.70%	0.79%	0.75%	0.34%	7.32%
	6	1.08%	0.68%	0.71%	0.47%	0.27%	0.36%	0.41%	0.49%	0.58%	0.56%	5.62%
	7	0.39%	0.63%	0.53%	0.42%	0.36%	0.28%	0.20%	0.28%	0.33%	0.53%	3.93%
	8	0.71%	0.26%	0.32%	0.23%	0.35%	0.40%	0.29%	0.23%	0.17%	0.27%	3.23%
	9	0.02%	0.02%	0.02%	0.02%	0.08%	0.17%	0.26%	0.35%	0.50%	0.50%	1.95%
Grand Total		11.57%	11.00%	11.05%	10.83%	9.20%	9.58%	9.59%	8.88%	8.74%	9.57%	100.00%

(a-4)

The above three tables shows significant trend of the higher the hazard rank, the higher the coming back to shop immidiate respose rate is. This pattern is consistent for first month after scoring, second momth after scoring and two month combined after scoring. It is also found that Hazard rank yeilds best time to purchase prediciotn for mid-high survival rank

How are we going to leaverage the finidngs and insights to improve targeting?

1. Enroll the time to response dimention and select names by hazard rank high to low if offer quantity is limited, i.e. offer quantity is less than one logistic rank. Instead of random selection, now we can leaverage time to purchase to optimiza offer results.
2. If larger scale selection, such as 40% of total poluation. We can use the red line on the bottom table as reference line, select down from the left side of table as those are names with higher instaniouse purchase likelhoood and replace those who are above red line on the right side of table as they have general high propensity to shop in the future 12 month but the event is rarely to happen at the time when the rencent event takes place.

		Hazard Rank										
		0	1	2	3	4	5	6	7	8	9	Grand Total
Prop Rank	0	1.99%	2.19%	2.16%	2.17%	1.90%	2.05%	2.09%	1.96%	2.33%	5.08%	23.94%
	1	1.89%	2.05%	2.01%	2.11%	1.80%	1.84%	1.84%	1.63%	1.58%	1.25%	18.00%
	2	1.79%	1.72%	1.83%	1.82%	1.50%	1.60%	1.61%	1.28%	1.01%	0.47%	14.63%
	3	1.49%	1.47%	1.50%	1.42%	1.30%	1.35%	1.24%	0.99%	0.81%	0.29%	11.87%
	4	1.22%	1.19%	1.11%	1.24%	1.02%	0.94%	0.95%	0.87%	0.68%	0.28%	9.51%
	5	0.99%	0.79%	0.86%	0.90%	0.61%	0.59%	0.70%	0.79%	0.75%	0.34%	7.32%
	6	1.08%	0.68%	0.71%	0.47%	0.27%	0.36%	0.41%	0.49%	0.58%	0.56%	5.62%
	7	0.39%	0.63%	0.53%	0.42%	0.36%	0.28%	0.20%	0.28%	0.33%	0.53%	3.93%
	8	0.71%	0.26%	0.32%	0.23%	0.35%	0.40%	0.29%	0.23%	0.17%	0.27%	3.23%
	9	0.02%	0.02%	0.02%	0.02%	0.08%	0.17%	0.26%	0.35%	0.50%	0.50%	1.95%
Grand Total		11.57%	11.00%	11.05%	10.83%	9.20%	9.58%	9.59%	8.88%	8.74%	9.57%	100.00%

(a-5)

The chart a-5 above also gives a exmaple that, when we select by two dimention model overlay, the lift for select 40% of total populaiton to target can yield at least 1.5% response rate.

Consider the 3.7% average in house response rate disclosued by DMA, 1.5% lift is very significant.

CONCLUSION

Layer Survival Prediction into Propensity prediction can complement with the time to event dimension which propensity model could not fully yield. The tool enables marketers to identify customers though in lower propensity but highly likely to make re-purchase shortly while deselect customers who though have high propensity to repurchase but the action may not occur for next marketing campaign. In such case, response can be lifted in general and camping selection can be done covering across more propensity diversity. The technique shared in this paper can be applied to different marketing area and even different industries. The most important take away is, to ensure optimal modeling results, most of the efforts have to be invested in model set up and data preparation.

REFERENCES

Allison, Paul D. 2003. Survival Analysis Using SAS®: A Practical Guide. Cary, NC: SAS Institute Inc.

Altman, E. I. 1968. Financial Ratios, Discriminant Analysis, and the Prediction of Corporate Bankruptcy. Journal of Finance 23: 589-609.

Junxiang Lu, 2013 Predicting Customer Churn in the Telecommunications Industry — An Application of Survival Analysis Modeling Using SAS, SUGI27

ACKNOWLEDGMENTS

I would like to express my great thankfulness to the whole Predictive Analytics team under Marketing Insights Department at Alliance Data Systems, Inc,, for providing the opportunity to work on this paper.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Hsin-Yi Wang

Enterprise: Alliance Data Systems, Inc.

Address: 3100 Easton Square Pl

City, State ZIP: Columbus, Ohio, 43219

E-mail: hsinyi.wang@alliancedata.com