Customer Analytics: Assignment-1

Due: 09.06.2020 23:59 PM

Part-I (25 pts)

Question-1 (5 pts): Describe the following terms:

Customer Equity:

Defection:

PCV:

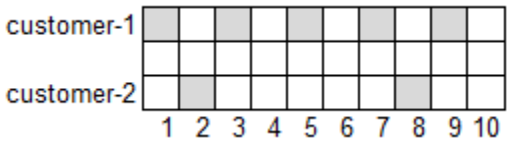
Right-censored:

Retention:

Question-2 (3 pts): Name a probability distribution that can be used to model customers inter purchase time behavior. Write down its probability density function.

Question-3 (2 pts): Plot, roughly, how retention rate of a customer group expected to evolve over time. Let y-axis shows the retention rate and x-axis the time.

Question-7 (6 pts): Assume that you would like to estimate the activeness level of the following customers for period t=11:



a) Which customer is more likely to churn (show calculations) at t = 11? (3 pts)

b) Assume that the time scale of the above graph is in months and each grey box shows customer purchase times. Further assume that each transaction has monetary value of $100. Write down the LTV formula for each customer explicitly. Use discount rate as 1% per month. (3 pts)

Question-8 (2 pts): If you would like to model customer's next time to purchase behavior, which of the following approach **cannot** be used?

a) Geometric distribution
b) Exponential distribution
c) Markov chains
d) Logistic regression
e) Classical multiple linear regression

Question-9 (2pts): Which of the following(s) are **false** for RFM analysis?

I – Each RFM cell needs to have equal amount of customers
II – The priority (the order of importance) is dependent on the context
III – The BEI value of cell 123 is greater than that of 321
IV – There might be cases where lift value for all cells are below average lift.

a) Only I
b) I and II
c) II and III
d) I, III and IV
e) I, II, III and IV

Question-10 (5 pts): Consider the following matrix, showing customers' likelihood to buy over time

		Brand purchased next time		
		A	B	C
Brand currently purchased	A	70%	20%	10%
	B	10%	80%	10%
	C	25%	15%	60%

Assuming that the customer has purchased brand A in t = 1. Answer the following (show calculations)

a) What is the probability that she won't purchase brand A in t = 2? (3 pts)

b) What is the probability that she will purchase brand A in t = 3? (2 pts)

Part-II (75 pts)

Question-1 (30 pts): Consider the brand switching example discussed in class. Assume that following information is given for three customers:

Customers	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
C1	1	2	1	1	1	2	1	2	1	1
C2	2	2	1	1	1	2	1	2	2	2
C3	1	2	2	1	1	2	2	1	2	2

In the above table, rows indicate customers and columns indicate the time periods (last 10 periods, assuming today is represented with t = 0, hence -1, -2, … -10 represent last ten purchases). Values in the table indicate the purchased brand ID. For example, Customer1 (C1) purchased brand 1 in period -10, then brand 2 in period -9, then brand 1 in period -8 etc.

a) Define your state space and **clearly** describe a procedure to estimate the transition probabilities from the provided data. (5 pts)
b) Estimate the transition probabilities using above data and provide the final transition matrix. (5 pts)
c) Now assume that customers switch between brands with the same probability, say *p*. Estimate *p* using maximum likelihood estimation. State your assumptions clearly and show all your work. (15 pts)

Question-3 (20pts):

a) Assume that we predict retention rates of a given cohort five years into the future. Predicted retention rates are listed in the below. Please fill out the missing values (following the columns headers) in the table. (5 pts)

Tenure	Predicted Retention Rate	Churn Rate	Survival Rate	Expected Number of Active Customers
1	30%			
2	50%			
3	65%			
4	70%			
5	75%			

b) Assume following are given:
- initial cohort is consisting of 1000 customers,
- tenure is given in years,
- customers purchase, on the average, goods that are worth $1000 per year,
- discount rate is 10% per year.

Calculate net present total customer equity of the firm. (10 pts)

Question-4 (25 pts):

Consider the data presented in Table 1.

Table 1 Market basket data.

Customer ID	Transcation ID	Items Bought
1	0001	{a,d,e}
1	0024	{a,b,c,e}
2	0012	{a,b,d,e}
2	0031	{a,c,d,e}
3	0015	{b,c,e}
3	0022	{b,d,e}
4	0029	{c,d}
4	0040	{a,b,c}
5	0033	{a,d,e}
5	0038	{a,b,e}

a) Compute the support for items {e}, {b,d}, and {b,d,e} by assuming each transaction ID as a market basket.

b) Use the results in part (a) and compute the confidence for rules {b,d} -> {e} and {e}->{b,d}. Is confidence a symmetric measure?

c) Recompute part a) by assuming each customer ID as a market basket. Each item should be assumed as a binary variable (1 if an item appears in at least one transaction of the customer, 0 otherwise)

d) Use the results in part c) and compute the confidence for rules b,d} -> {e} and {e}->{b,d}.

e) Compute the support, confidence and lift for the rule {a,b}->{e}.