Multi-state loan transition model

When constructing a “healthy” loan portfolio, probability of default (PD) is one of the primary factor to consider. PD describes the likelihood of default for a loan. In the long period, we want to estimate the probability of default by tracking the loan status’s movement at each time t>1 until the loan being paid off or charged-off. This is determined by multi-state loan transition model. In this write up, it provides a high level overview of probability of loan transition framework, together with its implication for Affirm.

**Methodology**

Before digging into multi-state loan transition framework, two concepts are described here which constitute the foundation of multi-state loan transition model: Markov model and multinomial logistic model.

Markov model

The assumption of Multi-state loan transition model is it follows Markov process. Markov model assumes future event depend only on the current state but not the past state. The model can be formalized as following:

Where represents the value at current state, and represents value at any state prior to .

Multinomial logistic regression

Multinomial logistic regression generalize the logistic model with two outcomes to more than two outcomes. A general form of multinomial logistic model is as follows:

……

where K represent outcome level and X represents a vector of covariates. β is the logodds when covariate increase by 1 unit.

Multi-state loan transition model

Putting Markov model and multinomial logistic regression model together, the multi-state loan transition model can be constructed to model Affirm loan portfolio. To start with, we defined the following state of a loan:

1. Current: all of the principal and interest that due being paid on time; there is no miss-payment.
2. Delinquent: one or more payments are missed. Delinquency is measured in days, such as 30 days delinquent, 60 days delinquent, .etc.
3. Charge off: the loan is in outstanding status, this often happens when loan hits 120/180 day past due.
4. Prepaid: the loan is prepaid before its official due date.

We want to model the transition path of a loan along the timeline until it is fully paid off or charged-off. Below is a full transition cycle for a loan starting from current. For simplicity, we consider a loan as charge-off once it passed 60 days delinquency.

Prepay

CurrentNNTNNT

Dq30

Dq60NNTNNT

ChgoffNNTNNT

In that, we want to track all the possible transition state for a loan. For example: a loan in current state can either stay in current, or transit into Dq30, or being prepaid. The table below provides the corresponding possible delinquency transitions from time t to t+1.

**Transition matrix**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Current | Dq30 | Dq60 | Dq90 | Dq120 | Charge-off | Prepaid |
| Current | current | C- Dq30 |  |  |  |  | C-Prepaid |
| Dq30 | Dq30-C | Dq30- Dq30 | Dq30- Dq60 |  |  |  | Dq30-Prepaid |
| Dq60 | Dq60-C | Dq60- Dq30 | Dq60- Dq60 | Dq60- Dq90 |  |  | Dq60- Prepaid |
| Dq90 | Dq90-C | Dq90- Dq30 | Dq90- Dq60 | Dq90- Dq90 |  |  | Dq90- Prepaid |
| Dq120 | Dq120-C | Dq120- Dq30 | Dq120- Dq60 | Dq120- Dq90 | Dq120- Dq120 | Dq120- Charge-off | Dq120- Prepaid |

For a current loan, it can either transition into Dq30, or pay the balance and remain in current, or prepays and close the account.

For the loan in Dq30 state, the loan can either cures by making two payment, or stays at Dq30 by making one payment, or transition to Dq60, or prepays.

The multi-state loan transition model is then model the transition probabilities from any time , to each of the next possible state at t+1. And it assumes Markov process holds, that the next state of loan only depends on current state regardless of previous state, which means

where S represent the loan state.

For each of current/delinquency loan at time t, a multinomial logistic regression (described in previous section), or tree based models are used to predict transition probabilities. A separate multinomial logistic model is fit for each starting point. Using the above transition matrix as an example one multinomial model will model the transition probability from current to (current, Dq30, Prepay) ; one will model the transition probability from Dq30 to (current, Dq30, Dq60, prepay), etc. Thus, there are 5 multinomial logistic regressions in total. When building classification models, loan attributes will also be added into model as covariates. The loan attributes including lender’s credit score, loan term length, and loan interest rate, which can be dynamic or static.

**Implications**

The multi-state loan transition model requires the payment behavior satisfy the Markov process. However, in reality, the state of a loan at t+1 may also correlated with the state at t-1. For example, at time t-1, the loan is in 90 days delinquency, and at time t, the loan goes back to current. To predict the loan state at time t+1, there is a higher chance that loan goes back to 30 days delinquency instead of stay current or being prepaid. This arises from the fact that at time t-1, the loan is already in deep delinquency status.

Although we see pitfalls of the model, it still provides a very nice framework to model the loan behaviors. The transition network covers all possible states, and to model the transition probability, various machine learning models can be used, from traditional logistic model to tree based method (random forest, gradient boosting trees). From testing prospective, it gives a high accuracy or area under ROC curve. The multi-state loan transition model can also be fast implemented.

Implementing multi-state loan transition model can help company (Affirm) (1) optimize the loan portfolio, reduce credit risk and (2) forecast loss in long term.

1. From credit risk prospective, with new applications, the default or loass can be easily estimated along the time line. That is, at each time t, we can estimate the transition probability of the loan into each of next possible state. Based on the estimated probability, we can further calculate the cash flow for each possible loan state. The process is iterated until the loan being paid off or charged-off. Lastly, the charge off amount can be obtained (if the loan finally transited into charge off state). The charge off amount can be considered as the final loss for each application. On one side, the application with negative loss or gain can have a higher chance of being approved; on the other side, for the applications with positive loss, a higher interest rate maybe desired to offset the potential positive loss.
2. From loss forecasting prospective, with existing applications, based on their current state, we can continue to forecast the next state until all loans being paid off or charged-off. We sum up all the amount at charge-off state and take it as the total loss for the existing portfolio, and that can help company manage the potential risk.