Multi-state Loan Transition Model

When constructing a “healthy” loan portfolio, probability of default (PD) is one of the primary factors to consider. PD describes the likelihood of default for a loan. In the long-term period, we want to estimate the probability of default by tracking the loan status’s movement at each time t>1 until the loan is paid off or charged off. This can be determined by a multi-state loan transition model. In this write up, it provides a high level overview of loan transition framework, and discusses how it would help to manage risk of the loan portfolio for Affirm.

1. **Methodology**

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

1.1 The Markov model

The assumption of the Multi-state loan transition model is that it follows the Markov process. The Markov model assumes the future event depends 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 .

1.2 Multinomial logistic regression

Multinomial logistic regression generalizes the logistic model with two or more outcomes. A general form of multinomial logistic model is as follows:

……

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

1.3 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 is 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. Charged off: the loan is in outstanding status, this often happens when the 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 charged off once it passed 60 days delinquency.

CurrentNNTNNT

Dq30

Dq60

ChgoffNNTNNT

Prepay

In that, we want to track all the possible transition states for a loan. For example: a loan in the current state can either stay in current, or transit into Dq30, or be prepaid. The table below provides the all the corresponding possible delinquency transitions from time t to t+1 (assume charged off after Dq120).

**Transition matrix**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Current | Dq30 | Dq60 | Dq90 | Dq120 | Charged 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- Charged off | Dq120- Prepaid |

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

For the loan in Dq30 state, the loan can either be cured by making two payments, 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 states at t+1. And it assumes Markov process hold that the next state of loan only depends on current state regardless of previous state, which means

where S represents the loan state.

For each of the current/delinquency loans 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 the model as covariates. The loan attributes include lender’s credit score, loan term length, and loan interest rate, which can be either dynamic or static.

1. **Implications**

The multi-state loan transition model requires the payment behavior to satisfy the Markov process. However, in reality, the state of a loan at t+1 may also correlate 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 staying 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 the traditional logistic model to tree based methods (random forest, gradient boosting trees). From testing perspective, it gives a high accuracy or area under ROC curve. The multi-state loan transition model can also be fast implemented.

Implementing a multi-state loan transition model can help company (Affirm) with:

1. Optimizing the loan portfolio and reduce credit risk. With new applications, the default or loss can be easily estimated along the timeline. That is, at each time t, we can estimate the transition probability of the loan into each of the next possible states. Based on the estimated probability, we can further calculate the cash flow for each possible state. The process is iterated until the loan is paid off or charged off. Lastly, the charged off amount can be obtained (if the loan finally transited into charged off state). The charged 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, an increasing interest rate may be desired to offset the potential positive loss.
2. Forecasting loss in the long term. With existing applications, based on their current state, we can continue to forecast the next state until all loans are paid off or charged off. We sum up all the amount at charged off state and take it as the total loss for the existing portfolio, and that can help the company to manage the potential risk.