Using the credit data set, coming from the European bank, we estimate the mover-stayer model, which is an extension of the Markov chain. This model assumes that the population is heterogeneous: there are “stayers” and “movers”. “Movers” evolve according to a Markov Chain with the one-step transition matrix, while “stayers” never leave their initial states. The probability of a customer being a stayer in a paid up state is modeled using the logistic regression.

We use the EM (expectation-maximization) algorithm to estimate the parameters of the mover-stayer model. It consists of two steps: the expectation step and the maximization step. In the E-step we compute the expected number of stayers in paid up state given the data and the values of the parameters. In the M-step we used the expected number of stayers from the E-step. The E-steps and M-steps were repeated until a convergence criterion was satisfied. The EM algorithm was built using seven covariates.

On the basis of the results received, we were able to distinguish covariate patterns for “stayers”, for “movers” and patterns that were shared among these two groups. We used bootstrap to estimate out of sample misclassification rate for the covariate patterns.

The proposed model may be customized for use by a lending institution (bank), which considers granting a loan to an applicant.