Churn Modeling

1. Defining the Time Window

The last date available is 02 February 2007. We will observe the dependent variable for a year.

End of dependent window = 31 December 2006 Start of dependent window = 01 January 2006 -------lag---------- End of independent window = 03 December 2005 Start of independent window =

1. Select qualified customers for the basetable

In this case, donors have to be active during the independent period. Given,

Exrelactcd – Activity code of the donor (FP means active) Extrelstdt – Start date of the relationship Exreldaten – End date of the relationship (Missing: not ended)

So, we define an active customer as one with - activity code = FP, - start date before end of IW - end date after start of DW [means he was active during IW] or missing [means still a donor],

1. Create Independent [PREDICTOR] Variables

Possible predictor variables for churn are

- Recency Frequency Monetary Value,
- Communication, Payment Information etc.

Note: Only use data from the independent window to create the predictor variables (Xs)

1. Create Dependent [PREDICTED] Variables

If the relationship ending date (exreldaten) is non-missing and lies within the DW, it means that account has churned.

1. Create the Basetable by merging data
2. Impute missings

For numeric - choose mean/median (check for skewness) For categorical - choose the mode of the distribution

1. Build Models! Find the best one on the Validation sample and Test its performance on the hold-out sample.