

Titolo tesi: Progettazione e sviluppo di un sistema di *Scoring* per la previsione dei clienti a rischio di abbandono – settore bancario.

Autore: Elisabetta Rapali

ABSTRACT

La Banca e Nunatac, azienda consulente, si occupano di *Predictive Customer Relationship Management*. Tra gli obiettivi di *P-Crm* vi è prevenire il rischio abbandono di singoli clienti. *Churn prevention scoring models* possono avere importanza strategica per la fidelizzazione dei clienti, per la competitività d'impresa. L'Istituto creditizio progetta e realizza iniziative di *caring* verso i clienti identificati tramite *Scoring System*. Negli anni i sistemi di *Scoring* vengono monitorati al fine di valutare i cambiamenti intervenuti nella popolazione di riferimento. Nel 2018 un nuovo modello accoglie questi cambiamenti della *Customer Base* della Banca.

Ogni modello statistico probabilistico per la previsione della propensione all'abbandono definisce unità statistica, orizzonte temporale, variabile target, popolazione d'analisi e i criteri di segmentazione di quest'ultima. Il nucleo del progetto qui condensato è la progettazione e lo sviluppo del nuovo modello per un uno dei segmenti clienti individuati da Nunatac. In fase di sviluppo è stata costruita un'apposita *Customer Table* a partire dalle tabelle indicatori del data mart aziendale, selezionato un campione di clienti, gestiti i valori mancanti delle variabili indipendenti esplicative. La variabile dipendente è dicotomica: il modello stima probabilità a posteriori che questa assuma valore 1 in un determinato intervallo temporale. È utilizzato software Sas Enterprise Miner.

Nunatac raffronta le performance di 28 modelli alternativi in competizione fra loro mediante *Cumulative Lift* calcolata nel *Validation* dataset. Tutte le performance

sono valutate per la Banca cliente nel primo decile dei risultati ordinati secondo la probabilità stimata decrescente. Il confronto si divide in quattro gruppi di algoritmi: regressione logistica *stepwise* (Garside 1965, Beale *et al.* 1967, Newton e Spurrel 1967); CART (Leo Breiman, Jerome Friedman, Charles J. Stone, R.A. Olshen, 1984, *Classification and Regression Trees*, Wadsworth In., Bertmont, CA.); Gradient Boosting ("*Greedy Function Approximation: A Gradient Boosting Machine*," and "*Stochastic Gradient Boosting*" by Jerome Friedman, 2001); Random Forest (Leo Breiman 2001).

Per le caratteristiche del segmento di clientela di riferimento, il modello più efficace è risultato essere uno del gruppo *Gradient Boosting* che è attualmente in produzione presso la Banca. I suoi risultati, monitorati periodicamente da Nunatac, sono mezzo per assegnare priorità ai singoli clienti a rischio abbandono e indirizzare loro attività di *marketing automation*.