

Master Universitario in  
“Analisi dati per la business intelligence e la data science”

A.A. 2019-2020

**Modelli di propensity: un’applicazione in ambito bancario**

Laurent Barmaverain

**Abstract**

Il presente lavoro di tesi ha come obiettivo lo sviluppo di *modelli di propensity* in ambito bancario. In particolare, è stata stimata la probabilità che un cliente della banca acquisti un prodotto assicurativo nei tre mesi successivi alla data di analisi. La probabilità stimata dal modello è interpretabile come la propensione del cliente all’acquisto del prodotto. La prima parte della tesi è dedicata alla descrizione teorica dei modelli utilizzati, quali regressione logistica, alberi decisionali, random forest e gradient boosting. La seconda parte descrive l’applicazione di questi modelli in un contesto reale e ripercorre il processo logico seguito durante l’analisi. Per eleggere il modello migliore ci si è chiesti quale dei quattro riuscisse ad individuare una parte di clienti, compresa tra il 5% e il 20% di tutti i clienti della banca, in cui la probabilità di osservare possibili sottoscrittori della polizza fosse molto più alta rispetto alla media del campione: la curva di lift cumulata è stata quindi utilizzata come misura principale per confrontare i risultati dei modelli. Dal confronto appare evidente che il gradient boosting ottiene i risultati migliori, con una lift cumulata sul primo ventile (5%) del campione di validazione pari a 11,87. Inoltre, le variabili più importanti utilizzate dal gradient boosting sono principalmente di natura economico-finanziaria. Questo rende il modello interpretabile e coerente con le logiche di business della banca. Infine, con lo scoring del modello out-of-sample e out-of-time si dimostra che i risultati ottenuti con il gradient boosting sono generalizzabili e utilizzabili in un processo di produzione.