BANK MARKETING ANALYSIS

“Marketing is the process by which companies create value for customers and build strong customer relationships to capture value from customers in return.”

***Kotler and Armstrong (2010).***

Il dataset utilizzato per l’analisi è stato importato da [Uci Machine Learning Repository](https://archive.ics.uci.edu/).

Il dataset contiene dati relative ad una campagna di marketing bancario e lo scopo dell’analisi è quello di trarre insights e raccomandazioni al fine di ottimizzare le future campagne di marketing e, in questo caso specifico, attrarre maggiori clienti ed aumentare il tasso di conversione relativo alla sottoscrizione di un conto deposito a termine[[1]](#footnote-1).

L’approccio, al fine di ottimizzare la campagna di marketing sulla scorta del nostro dataset, comprende i seguenti passaggi:

* Importazione del dataset ed analisi esplorativa sulle singole variabili e sulla relazione di ogni variabile con la variabile di risposta (ovvero il risultato della campagna di marketing);
* Pulizia e trasformazione dei dati: rimozione di colonne irrilevanti, gestione dei valori mancanti, gestione degli outliers, trasformazione delle colonne categoriche in variabili numeriche (One-Hot-Encoding) e delle variabili categoriche binarie (“yes” e “no”) in colonne che contengano solo valori booleani;
* Utilizzare tecniche di Machine Learning per prevedere il risultato della campagna di marketing e per identificare quali sono i fattori che influenzano maggiormente il successo della campagna stessa (Feature Importance).

Nello specifico, le tecniche di ML utilizzate sono state le seguenti:

* Logistic Regression
* Decision tree
* Random Forest
* XGBoost

ANALISI ESPLORATIVA

Il dataset

note

campaign mettere 10 negli outliers

fare lo stesso con previous

creare una colonna dummies outliers con 1 o 0 se quella riga era un outlier

nei boxplot per la rappresentazione bivariata con var numeriche fai log su balance

la metrica in un dataset sbilanciato è l’f1 score, poiché l’accuracy non va bene, allora mostro l’f1 score

poutcome importance

month = controllare correlazione tra month e deposit per validare la feature importance

1. Il conto deposito è un prodotto che consente di ricevere una [remunerazione garantita](https://blog.moneyfarm.com/it/finanza-personale/ti-conviene-veramente-il-conto-deposito-analisi-dei-rendimenti/) sul capitale in giacenza con una scadenza massima di 5 anni. [↑](#footnote-ref-1)