



Progetto Digital Marketing

Sara Filip

Matricola: 852864

Fasi del progetto

PREMESSA

Analisi preliminari

Domande di business



CONFRONTO

Descrizione soluzione

Metodologia



SOLUZIONE

Evidenze

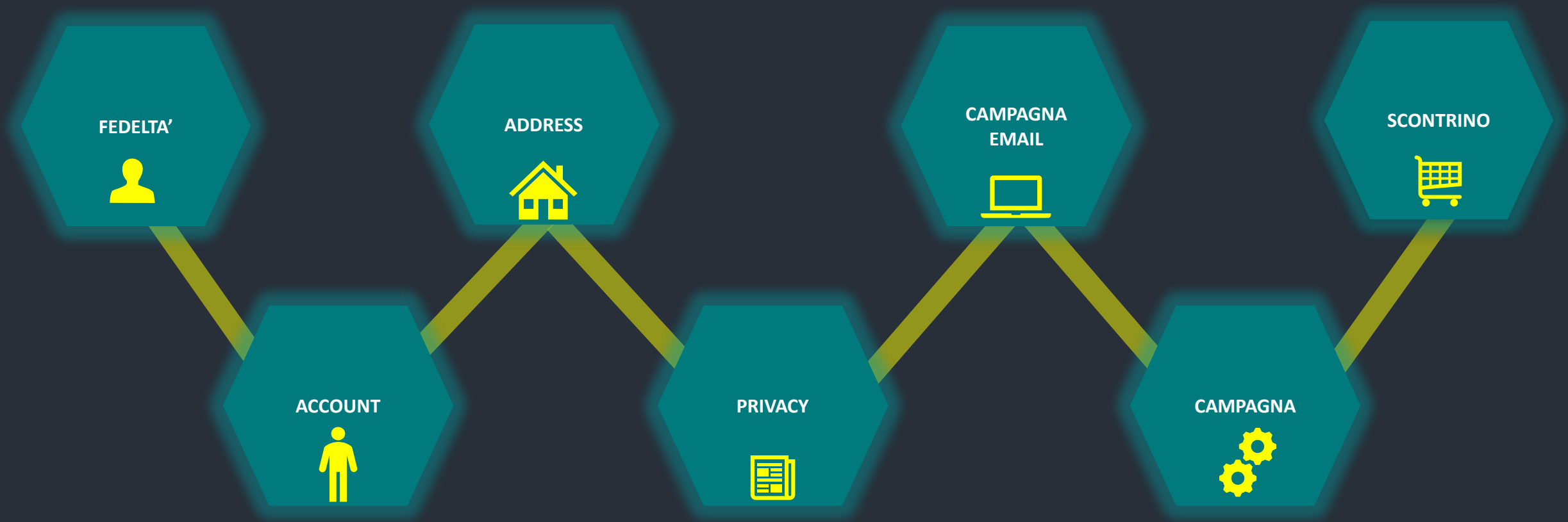
Conclusioni

Contesto

Si dispone di 7 dataset contenenti i dati relativi ad una società di vendita al dettaglio.

Lo scopo del progetto è presentare una strategia di comunicazione data-driven basata sui modelli di Machine Learning.

Dataset



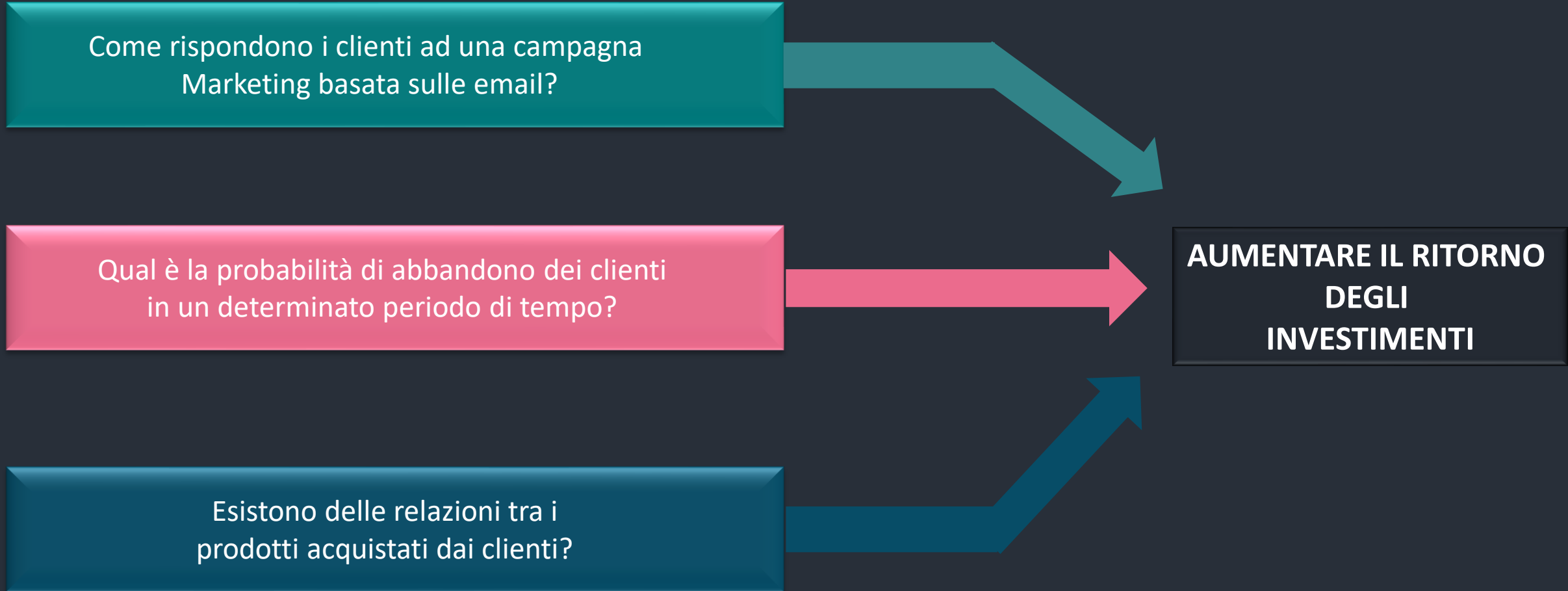
Domande di business

Come rispondono i clienti ad una campagna
Marketing basata sulle email?

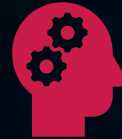
Qual è la probabilità di abbandono dei clienti
in un determinato periodo di tempo?

Esistono delle relazioni tra i
prodotti acquistati dai clienti?

**AUMENTARE IL RITORNO
DEGLI
INVESTIMENTI**



Soluzioni in risposta alle domande di business



Si sviluppa il modello RFM per decidere come differenziare le azioni di marketing tra i clienti



Si sviluppa il modello di Churn per determinare la probabilità di abbandono dei clienti. I dati possono essere utili a capire quando effettuare azioni di marketing ad hoc per prevenire l'abbandono dei clienti



Si propone la Basket Market Analysis per evidenziare eventuali relazioni tra i prodotti acquistati

Metodologia

Data cleaning

Pulizia dei dati, gestione missing value, rimozione record, definizione formato attributi

Preprocessing

Riorganizzazione dei dati in forma idonea alla modellazione (creazione nuovi attributi, operazioni di merge)

Data exploration

Analisi descrittiva degli attributi d'interesse e analisi quantitativa per la selezione della variabili da inserire nel modello

Modelling

Sviluppo del modello RFM, modello Churn e Market Basket Analysis

Modello RFM



Nel modello RFM
sono studiate 3
dimensioni



- **Recency**: quando il
cliente ha effettuato
l'ultimo acquisto



- **Frequency**:
frequenza di acquisto



- **Monetary Value**:
quanto il cliente
spende



Divisione della base
clienti in attivi e
inattivi



Focus on
clienti attivi

Evidenze del modello RFM

Il maggior numero di clienti appartiene al segmento *Bronze* (25932), seguito subito dopo dal segmento *Tin* (24967); *Gold* risulta invece il segmento con il minor numero di clienti (5586).

Si propone:

- di effettuare più azioni di marketing verso il segmento *Gold*, al fine di portare tali clienti verso il segmento *Diamond*
- mettere in atto azioni di marketing rivolte ai clienti di basso valore (*Cheap*) per portarli al segmento *Copper*
- individuare azioni di marketing specifiche per trasformare i clienti *Bronze* in *Silver* e i clienti *Silver* in *Gold*

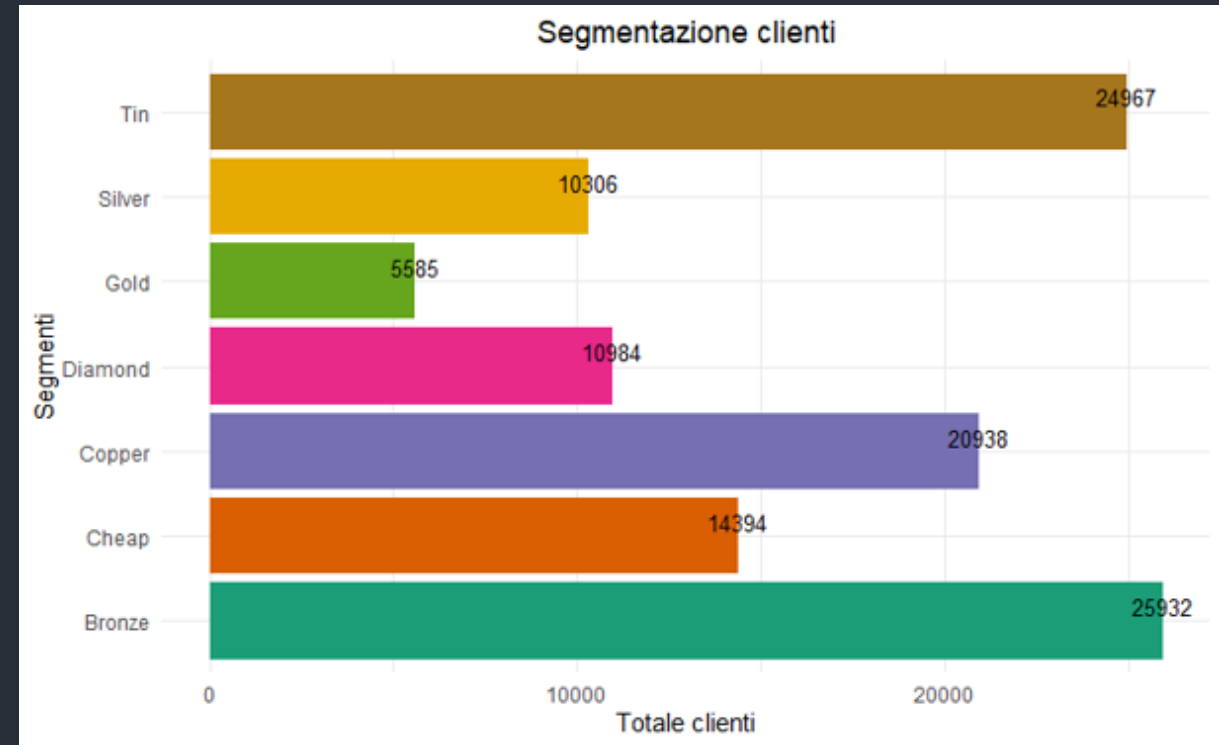


Figura 1: Segmentazione clientela

Evidenze del modello di Churn

- Circa l'80% dei clienti riacquista entro 60 giorni
- Periodo di studio: 1 Ottobre 2018 – 1 Gennaio 2019
- *Holdout period*: 1 Gennaio 2019 - 28 Febbraio 2019. All'interno di questo periodo si cercano i *churner*.
- Creazione variabile target: 35373 clienti non *churner*, 59752 *churner*
- Partizione dei dati in training e test set.

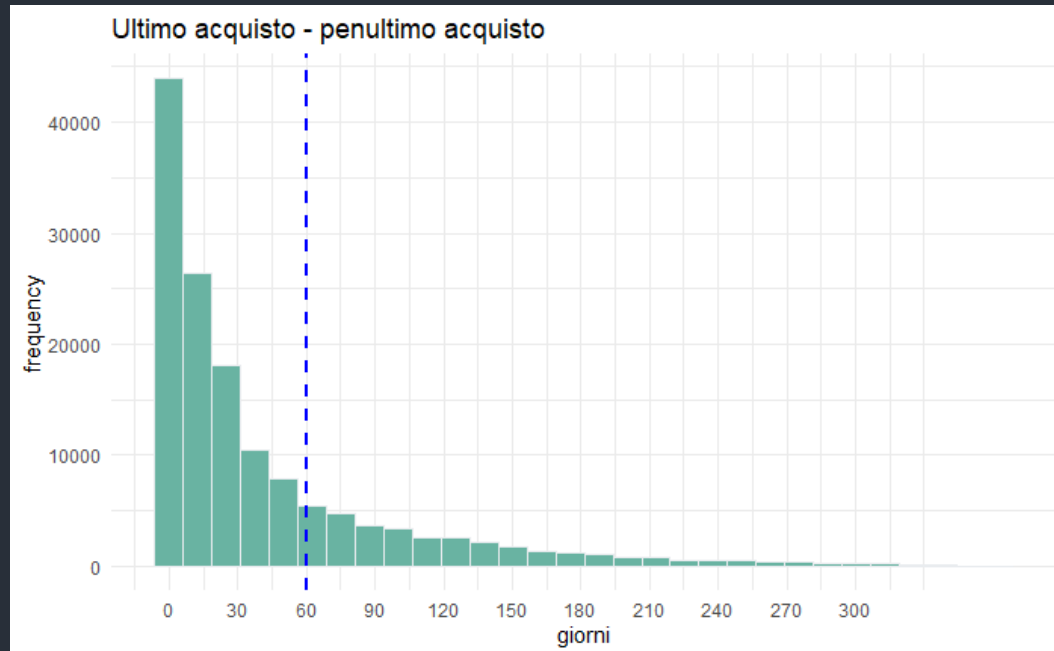


Figura 2: Giorni trascorsi tra penultimo e ultimo acquisto

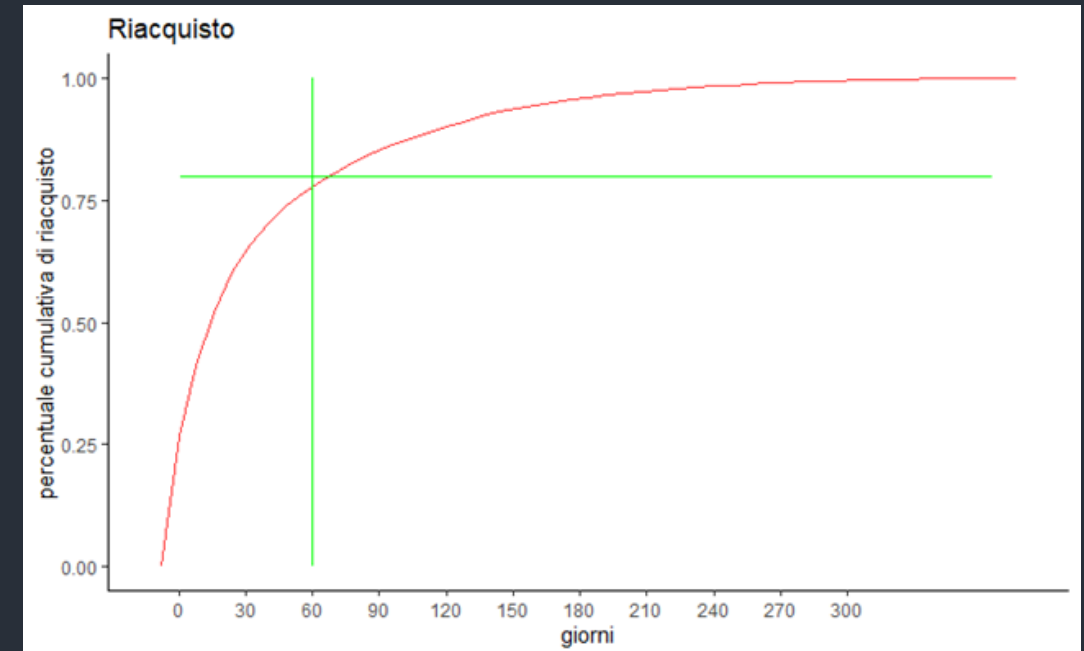


Figura 3: Circa l'80% dei clienti riacquista ogni 60 giorni

Evidenze del modello di Churn

Predittori: recency, spesa, totale acquisti, tipo di lavoro, ultimo stato di fedeltà, regione.

Modelli usati per prevedere la probabilità di abbandono dei clienti:

1. Modello di regressione logistica
2. Decision Tree
3. Random Forest

La scelta del modello di previsione può ricadere sul modello di regressione logistica o sul random forest; è meno consigliato l'utilizzo del decision tree perché ha un valore di accuracy minore, anche se di poco.

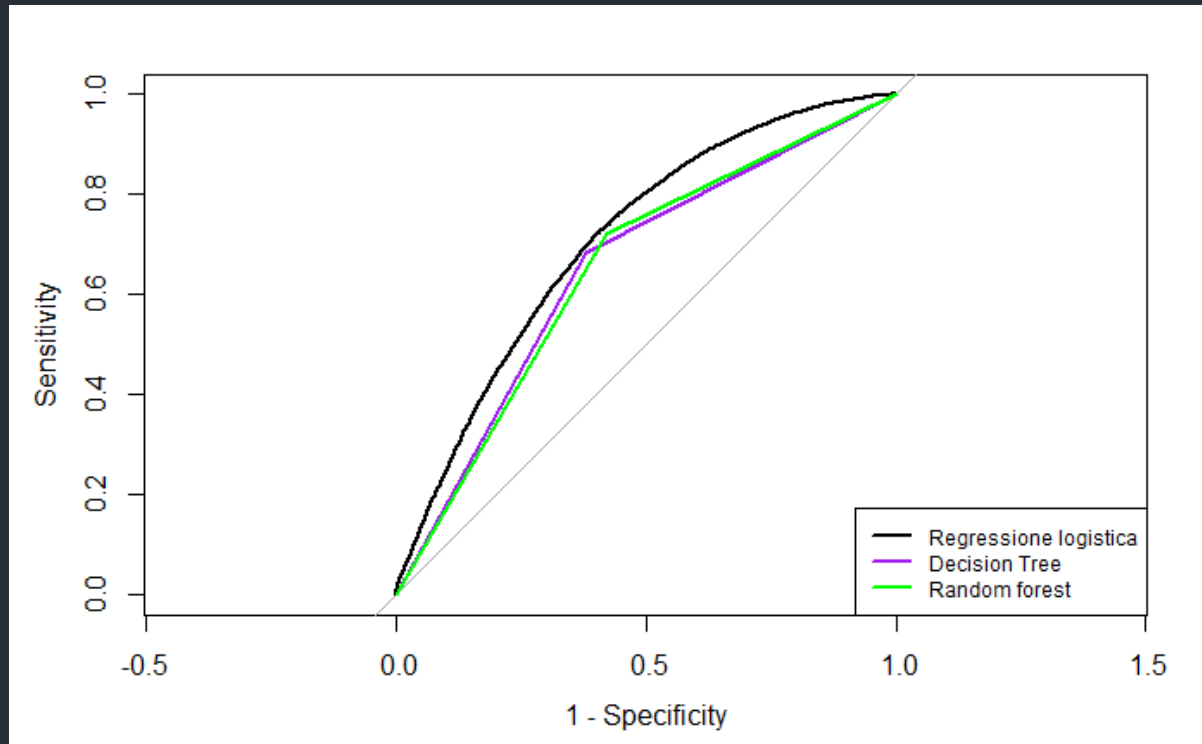


Figura 4: Curva Roc per i tre modelli

	Accuracy
Regressione logistica	0.66
Decision tree	0.65
Random forest	0.66

Tabella 1: Valori di Accuracy per i tre modelli

Evidenze della Market Basket Analysis

- Trasformazioni dei dati in transazioni
- Il prodotto più acquistato è il 33700716 (Figura 5)
- Individuazione di 47 regole di associazione tramite l'algoritmo «Apriori». Focus sulle prime 10 (Figura 6)

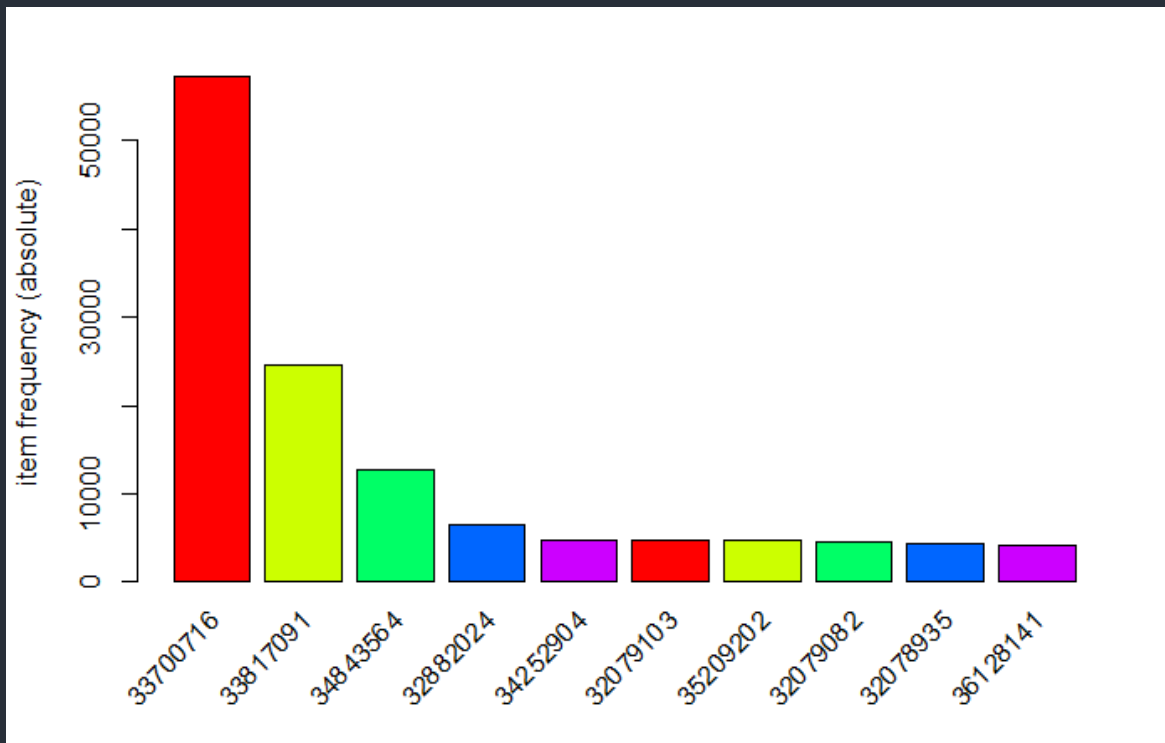


Figura 5: Prodotti più acquistati

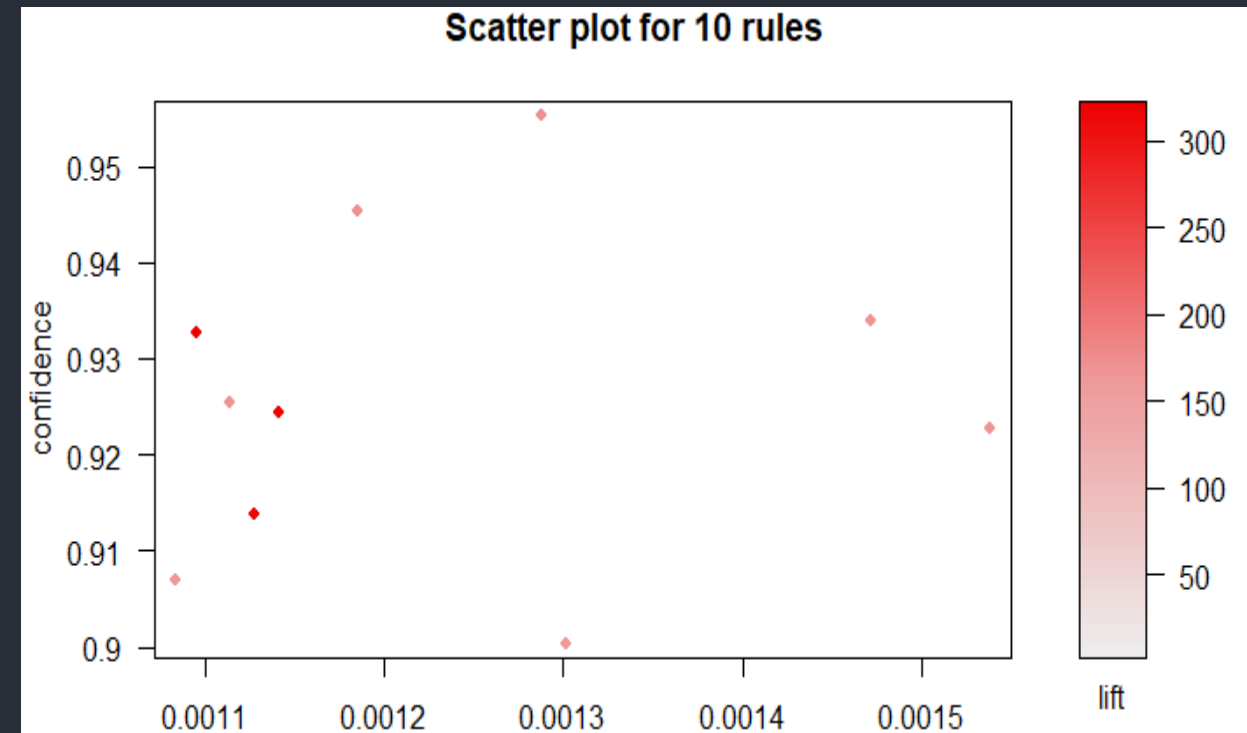


Figura 6: Prime 10 regole di associazione dei prodotti

Evidenze della Market Basket Analysis

Le regole di associazione evidenziano per esempio che:

- chi compra i prodotti {32078795,32079082,32842551} compra anche il prodotto {32079103} il 95.5% delle volte; questi prodotti sono stati acquistati insieme 1090 volte.

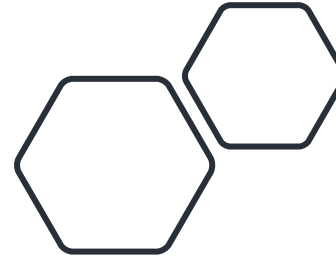
Prodotti acquistati	Acquisto associato	Livello di confidenza	Numero di volte
{32078795,32079082,32842551}	{32079103}	95.5%	1090
{32078935,32079082,32842551}	{32079103}	94.5%	1003
{32078795,32842551}	{32079103}	93.4%	1246
{36298381,36298416}	{36298353}	93.2%	927
{32078795,32078970,32079082}	{32079103}	92.5%	943



Conclusioni: azioni data-driven

- Adottare un approccio multi-canale: favorire il contatto con il cliente attraverso canali diversi per aumentare l'interazione
- Creazione di nuovi *touchpoint* per l'incontro di nuovi clienti e per l'ascolto delle esigenze del cliente; in questo frangente il cliente può diventare co-produttore di prodotti/servizi
- Effettuare azioni di marketing ad hoc per i diversi segmenti della clientela al fine di aumentarne la fidelizzazione
- Prevenire l'abbandono dei clienti a rischio attraverso azioni di marketing ad hoc (ad esempio offrire una maggior personalizzazione del prodotto/servizio erogato o in alternativa una promozione)
- Riorganizzazione dei prodotti in base alle relazioni emerse dalla Market Basket Analysis per aumentare la possibilità di vendita

Grazie per l'attenzione



Sara Filip

Matricola: 852864

E-mail:

s.filip@campus.unimib.it