

DIGITAL MARKETING

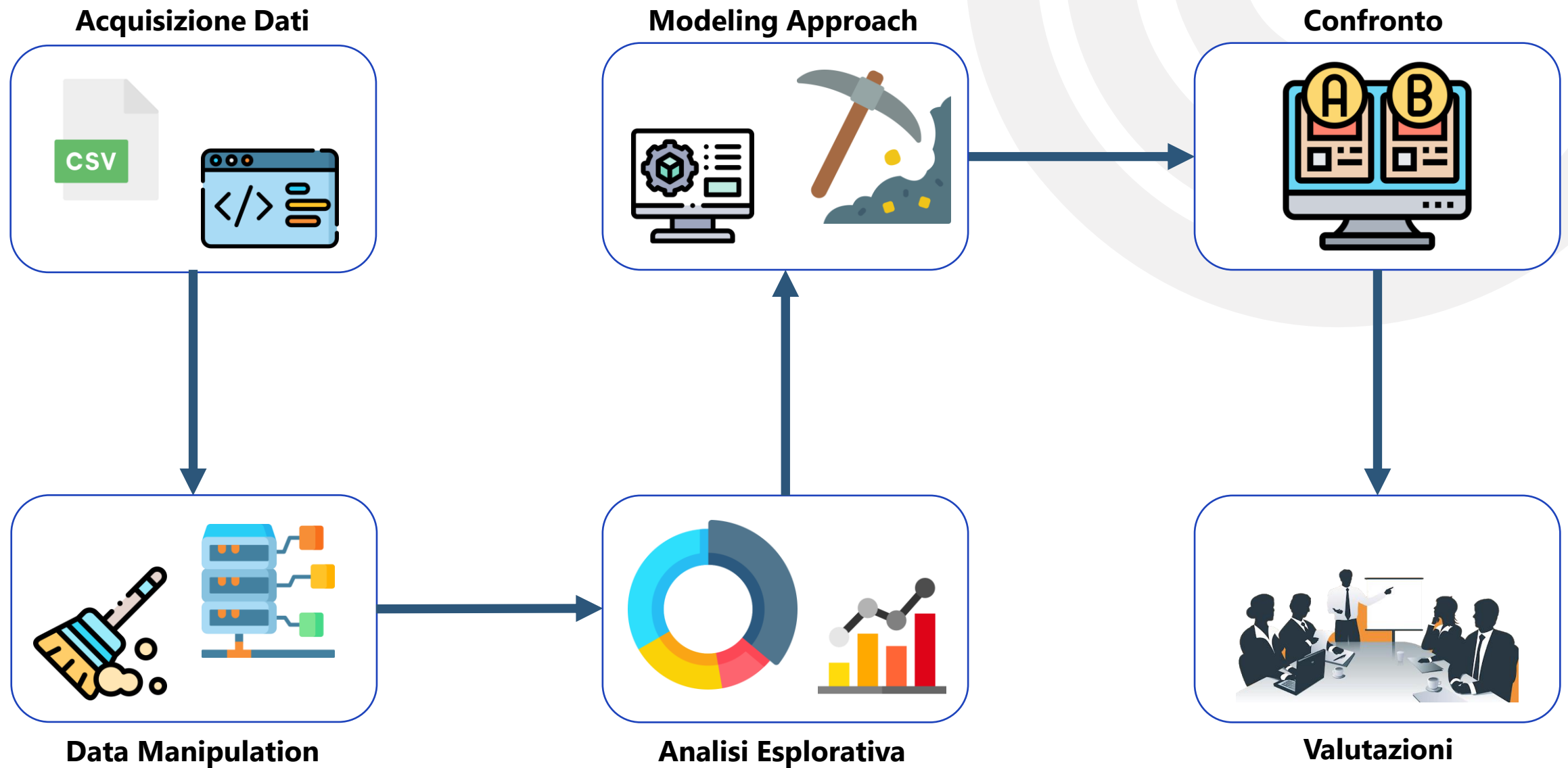
ALBERTO FILOSA



DIGITAL MARKETING

A.A. 2019 - 2020

PIPELINE



DATASET



FIDELITY

Cliente
Negozio
Fedeltà
Account Principale
Data Attivazione



ACCOUNT

Email
Cellulare
Tipo di Account
Tipo di Lavoro



ADDRESS

CAP
Provincia
Regione



PRIVACY

Flag Privacy
Flag Marketing



CAMPAIGN

Tipo di Campagna
Consegna
Tipo di Evento



TICKETS

Transazione
Importo
Sconto
Acquisto o Reso

BUSINESS QUESTIONS

Lo studio si basa su un approccio di comunicazione data – driven tramite costruzione di modelli di Machine Learning.

Sono stati utilizzati i seguenti modelli:



RFM

Recency, Frequency e Monetary sono delle misure di analisi di Marketing utilizzate per identificare i migliori clienti. Il modello RFM si basa su 3 principali fattori:

- Recency*: quanto tempo fa il cliente ha acquistato un prodotto;
- Frequency*: il numero totale di acquisti;
- Monetary*: Spesa totale degli acquisti.¹



MARKET BASKET ANALYSIS

La *Market Basket Analysis* è una delle principali tecniche utilizzate dalle aziende per individuare dei pattern nell'acquisto dei prodotti.

L'obiettivo dell'analisi è individuare la probabilità di identificare l'associazione tra i 100 prodotti più venduti.



PROPENSITY TO CHURN

Il *Churn* è un modello di Marketing che applicato ad una base di clienti ed identifica se uno di essi non acquisterà più un prodotto dato un certo periodo.

Esso è strettamente legato al profitto della azienda. Inoltre, identifica le aree dove il servizio clienti è meno presente.¹

ANALISI ESPLORATIVA

Per *Analisi Esplorativa* si intende il processo preliminare di analisi delle variabili tramite l'uso di statistiche generali come la media, mediana, moda e le distribuzioni di esse. Per la visualizzazione grafica è possibile costruire sia istogrammi che *boxplot*.

Prima della costruzione di modelli statistici, è necessario compiere questa analisi.

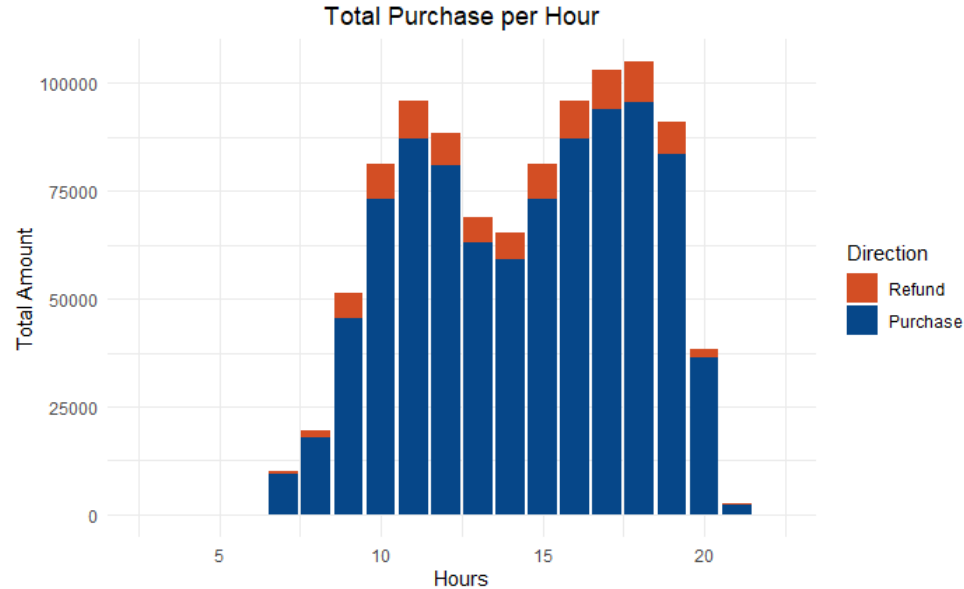


Figura 1. Distribuzione degli Acquisti per Ora

Nel primo è possibile notare la distribuzione degli acquisti (colorato in **blu**) di tutti i prodotti per ogni ora. Si osserva che negli orari di apertura e chiusura i clienti acquistano di meno, mentre gli orari di maggior frequenza sono le ore prima della pausa pranzo (dalle 10 alle 13) e subito dopo la giornata lavorativa (dalle 17 alle 19).

Nel secondo grafico è possibile notare la media giornaliera di apertura delle mail delle proposte di acquisto; circa il 70% dei clienti apre la mail il giorno stesso, mentre la percentuale raggiunge oltre il 90% dopo 4 giorni.

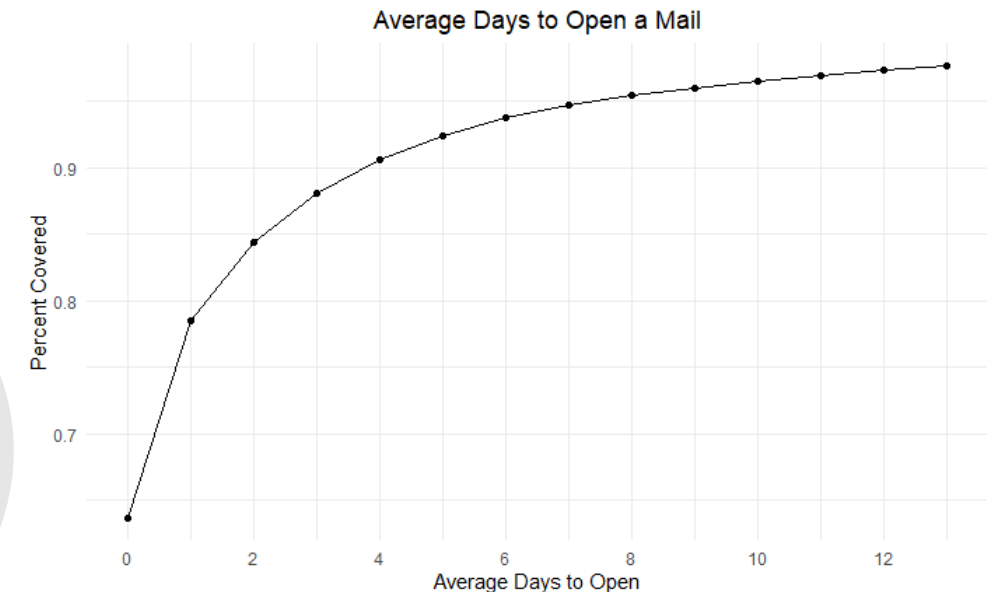


Figura 2. Distribuzione dei Giorni di Apertura delle Mail

MODELING - RF

I clienti più validi sono considerati in base a quanto recente hanno acquistato un prodotto, quanto spesso acquistano ed al valore totale speso. In questo modo è possibile raggruppare i clienti in categorie per valutare il valore dello stesso. Ad esempio, un cliente *Diamond* è molto più fidelizzato rispetto ad un cliente *Bronze*.

Dopo aver diviso la clientela in *Attivi* e *Non Attivi*, la si divide in gruppi in base alla distribuzione percentuale dei clienti per facilitare l'interpretabilità dei dati. La tabella mostra la divisione in percentuali:

Valore	Recency	Frequency	Monetary
Low	< 25%	< 2 Acquisti	< 25%
Medium	25% < X < 75%	2 < Acquisti < 5	25% < X < 75%
High	> 75%	> 5 Acquisti	> 75%

Tabella 1. Divisione Percentuale per costruzione Matrice RFM

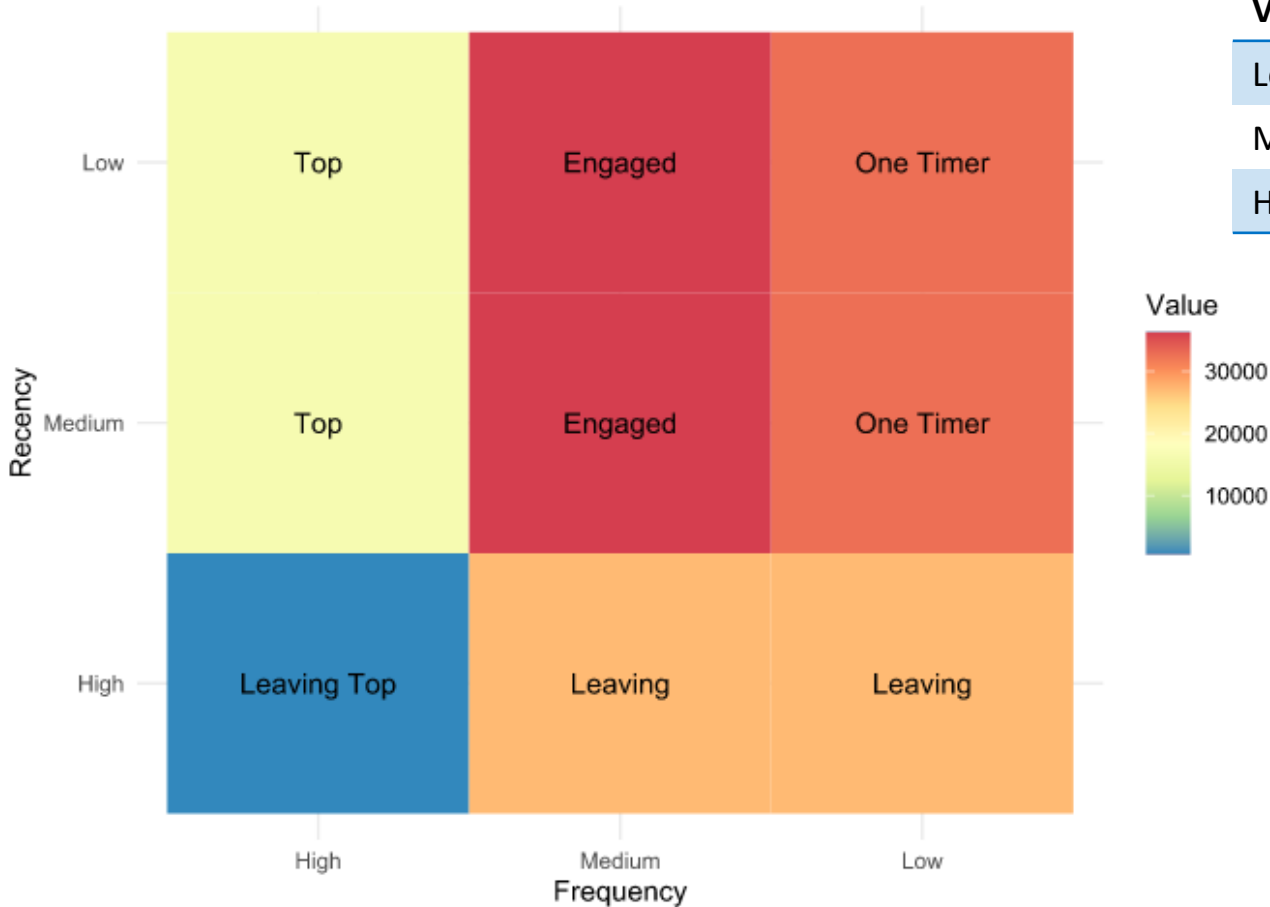


Figura 3. Matrice RF

In *Figura 3* è possibile osservare che la maggior parte dei clienti appartiene alla classe *Engaged*, mentre fortunatamente quella con il più basso valore è *Leaving Top*.

È necessario interagire maggiormente con gli *One Timer* in modo tale da trasformarli potenzialmente in *Engaged* e proporre degli sconti ai *Leaving* in modo da non lasciare l'azienda.

MODELING - RFM

Infine, si combina la matrice RF con i gruppi della *Monetary* per costruire la matrice RFM, con 7 diverse classi. Così facendo, è possibile differenziare le azioni di marketing lungo la clientela. Ad esempio, è possibile prendersi maggiormente cura delle persone raggruppate nelle più alte categorie, come i *Diamond* ed i *Silver*, mentre investire meno nelle persone di basso valore, come i *Tin* ed i *Cheap*.

Il gruppo con più clienti è il *Bronze*, con circa 25'000 persone al suo interno, mentre il gruppo *Gold* è quello con il minor numero. È necessario prendersi più cura delle ultime in modo da trasformarle in *Diamond*.

Come detto in precedenza, bisogna interagire maggiormente con il gruppo *Cheap*, che spende ed acquista poche volte, in modo da trasformarli in *Copper* od addirittura in *Bronze*.



Figura 4. Matrice RFM

MODELING - CHURN

Il modello di *Churn* prevede una stima della verosimiglianza che un cliente non acquisti più un prodotto della azienda costruendo un *propensity supervised model*. Per costruire il modello è necessario seguire questi passaggi:

1. Le variabili esplicative devono essere relative alla clientela, il tipo di iscrizione e le interazioni con l'azienda;
2. Controllo del *Churn*;
3. Stagionalità;
4. Escludere già i clienti che hanno lasciato l'azienda;
5. Outliers.

Si è scelto di assegnare come periodo di studio tra il 1° Ottobre 2018 ed il 1° Gennaio 2019, mentre il periodo di *Holdout* dopo la data di riferimento tutto il mese di febbraio.

Si assegna la variabile *Churn* con valore 1 se un cliente non ha acquistato un prodotto durante il periodo di *Holdout*, altrimenti 0.

No Churn - 0	Churn - 1
35'373	59'752

Tabella 2. Divisione Churn

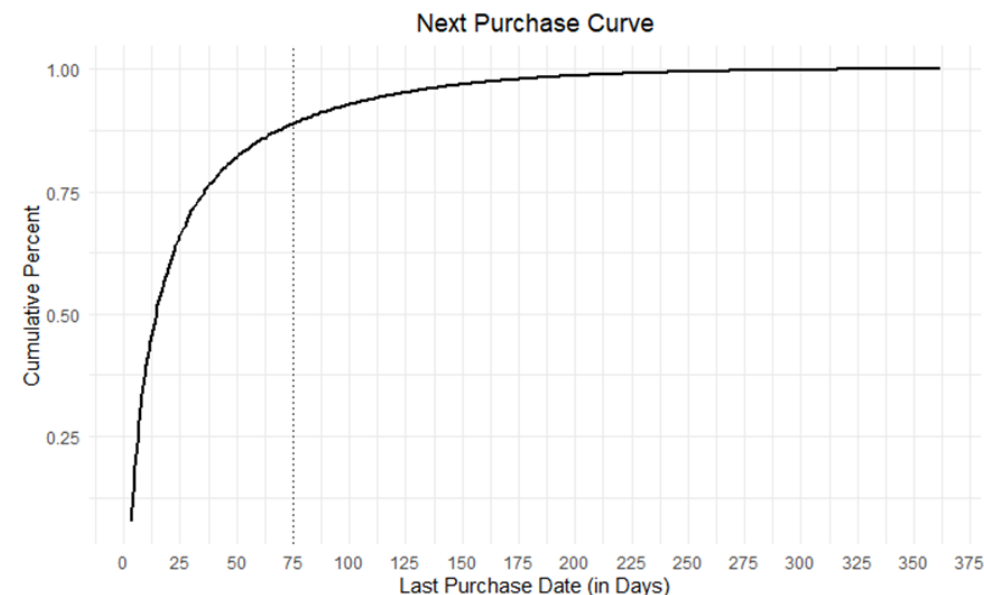


Figura 5. Next Purchase Curve

Si osserva che più della metà delle persone hanno acquistato almeno una volta nei tre mesi di studio.

La tabella non può essere relazionata con la *Next Purchase Curve* in quanto nella prima il cliente è considerato una volta sola, mentre nella seconda più volte.

MODELING - CHURN

Il dataset viene diviso in *Train* (70 %), in modo da allenare il modello per poi performare nei dati del *Test* (30 %) tramite la validazione *Holdout*. Le variabili esplicative utilizzati per la costruzione dei modelli di *Machine Learning* sono:

1. Recency;
2. Monetary;
3. Spesa Totale;
4. Regione;
5. Ultimo Acquisto;
6. Tipo di Lavoro.

Il modelli valutati per prevedere il Churn di un cliente sono:

1. Random Forest;
2. Regressione Logistica;
3. Recursive Partitioning and Regression Tree;
4. Lasso.

Il grafico Lift rappresenta la percentuale di persone classificate positivamente in base alla grandezza del campione.

Il modello vincente è la Regressione Logistica, con una Accuracy pari al 69.12 %. Dal grafico è possibile notare che considerando il 60 % dell'intero dataset si ha il 75 % di predire correttamente la classe.

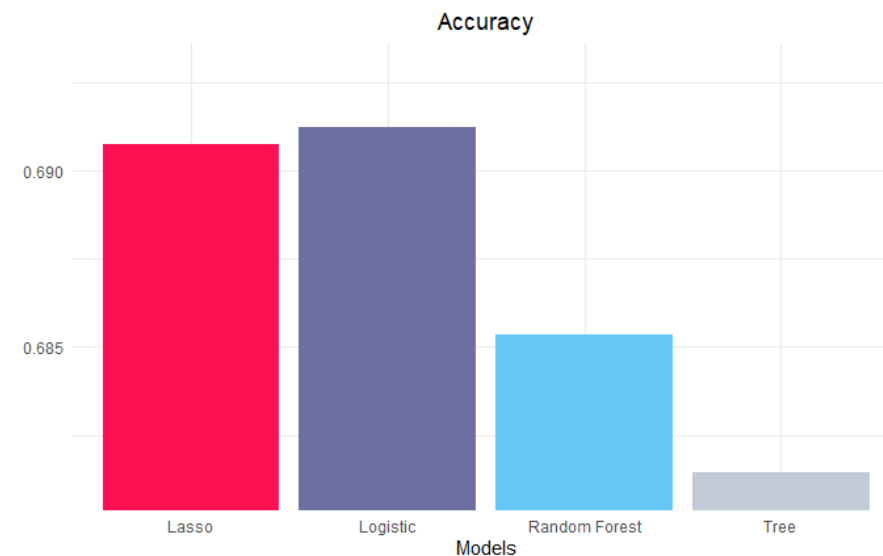


Figura 6. Accuracy dei Modelli

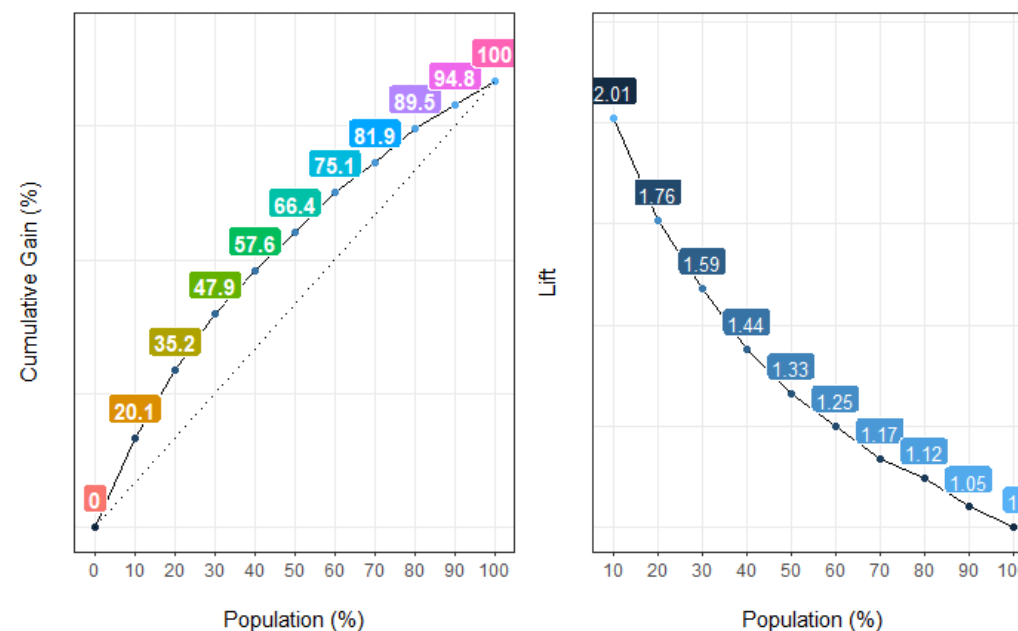


Figura 7. Gain e Lift del modello Logistico

MODELING – MARKET BASKET ANALYSIS

L'Association Rule è molto utilizzata per analizzare gli acquisti dei clienti o nei dati delle transazioni, in modo da identificare forte relazioni utilizzando misure di interesse.

Il prodotto nettamente più venduto è il 33700716, con circa 60'000 unità vendute; a distanza segue il prodotto 33817091, con oltre 20'000 unità acquistate.

Si osserva inoltre che i clienti acquistano maggiormente dai 3 ai 7 prodotti per ogni transazione, con un valore massimo di 212219 per 3 prodotti. In media i clienti hanno acquistato 6 prodotti, con un valore massimo di 147 pezzi.

Si costruisce l'Association Rule con supporto almeno dello 0.1 % ed una confidenza dello 80 %. Si osserva che il 95 % dei clienti che acquista i prodotti {32078795, 32079082, 32842551} acquistano anche il prodotto 32079103 con una alta relazione tra i prodotti.

Il grafico sottostante mostra il supporto in relazione della confidenza delle migliori associazioni tra prodotti:

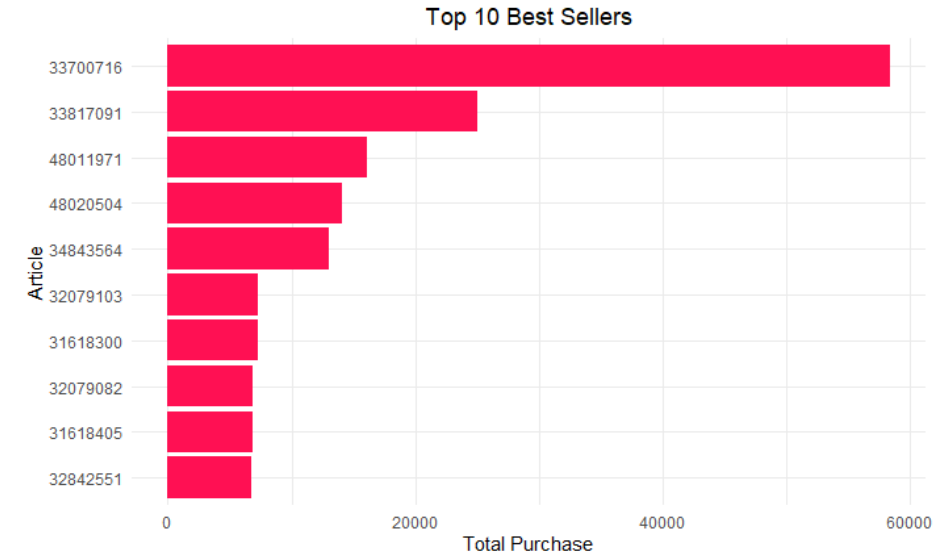


Figura 8. Prodotti più Venduti

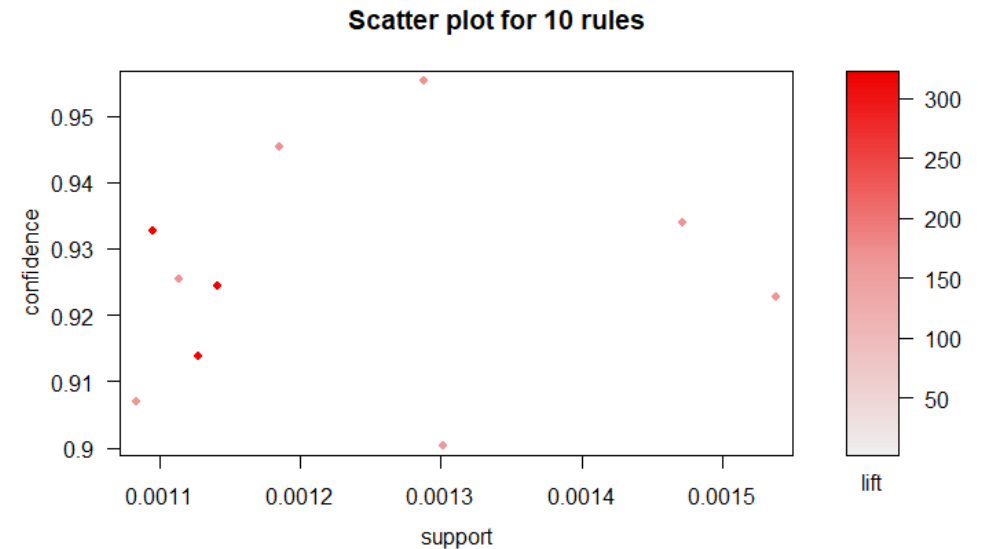


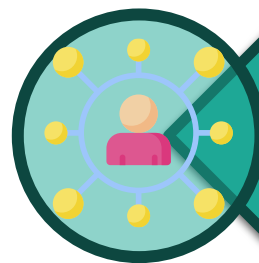
Figura 9. Migliori Associazioni di Prodotti

AZIONI DATA - DRIVEN



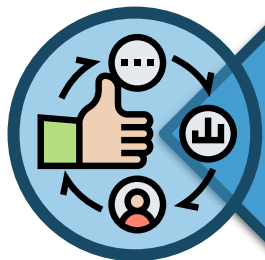
Campagne Promozionali

Segmentare le promozioni dai profili dei cliente tramite RFM



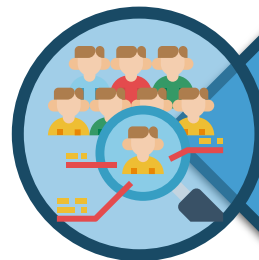
RFM

Interagire di più con i profili con pochi acquisti proponendo sconti su determinati prodotti



Strategia a Più Canali

Comunicare ai clienti su diversi canali per migliorarne l'interazione



Churn

Fidelizzare maggiormente i potenziali Churner comprendendo i bisogni dei clienti



Digital Shopping

Sviluppare un sito di vendita Online



Market Basket Analysis

Riorganizzazione dei prodotti in base alle associazioni della MBA



GRAZIE



A.FILOSA1@CAMPUS.UNIMIB.IT



MATRICOLA - 815589