



Gruppi di lavoro

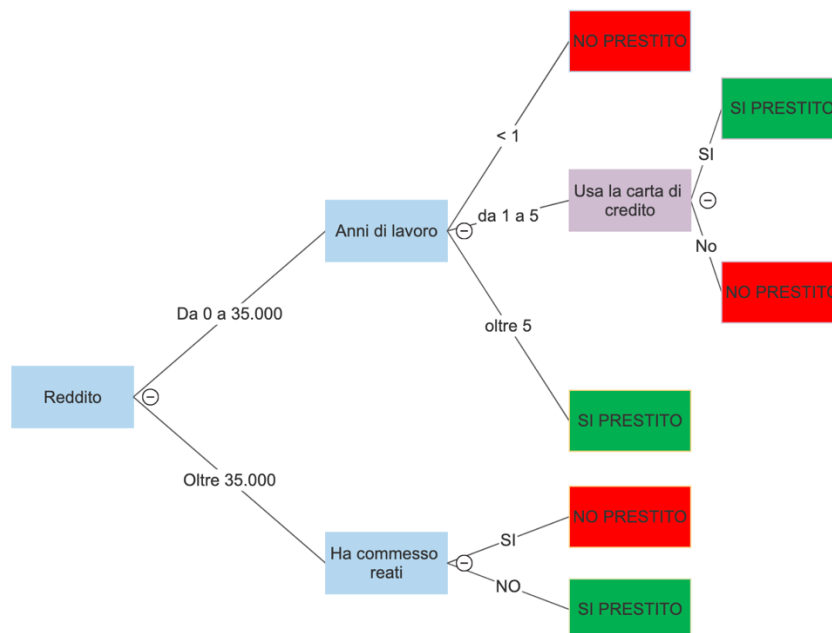
Nome	Gruppo di lavoro
Alessandro Sposetti	Finanza
Gabriel Radu Octavian	Finanza
Manila Ioffredi	Finanza
Maria Teresa Giosy Canto	Finanza
Angela Fortino	Logistica
Elisa Achilli	Logistica
Flavio Onori	Logistica
Francesca Prezioso	Logistica
Caterina Maria Malaspina	Marketing
Gioele Gueli	Marketing
Nicola Elia de Petris	Marketing
Antonio Andreacchio	Sanità
Chiara Bellanca	Sanità
Davide Lupattelli	Sanità

Gruppo di Lavoro – Finanza

Coding

Siete i data analyst di una startup fintech che permette ai clienti di un ecommerce di rateizzare il pagamento degli acquisti. Il reparto risk management ha sviluppato un decision tree basato sulle caratteristiche degli acquirenti: reddito annuale, reati, anni di lavoro, utilizzo di carta di credito. Se il cliente rientra nei parametri può richiedere il prestito altrimenti l'opzione gli viene negata, riducendo il rischio per la startup di prestare soldi a chi non potrà restituirli.

Create un algoritmo che restituisca correttamente il risultato del decision tree.



Big data

Nel settore fintech, l'analisi dei dati è fondamentale per identificare i clienti a rischio di insolvenza e per sviluppare strategie efficaci di gestione del rischio. Attraverso l'analisi dei dati, le aziende possono prevedere comportamenti finanziari, personalizzare le offerte e migliorare la gestione del credito. Ad esempio, molte istituzioni finanziarie utilizzano modelli predittivi basati su dati demografici e comportamentali per valutare la probabilità di default dei clienti.

In qualità di analisti nel reparto Business Intelligence di una startup fintech, il vostro compito è esaminare i dati dei clienti per identificare i fattori che contribuiscono alle difficoltà di pagamento. L'obiettivo è sviluppare modelli predittivi che consentano di anticipare il rischio di default e implementare misure preventive. Questo approccio



proattivo è essenziale per mantenere la stabilità finanziaria dell'azienda e offrire soluzioni adeguate ai clienti. Attraverso l'analisi dei dati, potrete individuare pattern e tendenze che influenzano la solvibilità dei clienti, permettendo all'azienda di adottare strategie mirate per minimizzare il rischio di default.

Obiettivo:

Identificare i fattori chiave che influenzano il rischio di default dei clienti, al fine di implementare strategie efficaci per minimizzare tale rischio.

Informazioni disponibili:

- **default:** 1 = cliente con difficoltà di pagamento (ha avuto ritardi nei pagamenti), 0 = altrimenti
- **CODE_GENDER:** Sesso
- **FLAG_OWN_CAR:** Possesso di autovettura
- **FLAG_OWN_REALTY:** Possesso di casa
- **CNT_CHILDREN:** Numero di figli
- **AMT_INCOME_TOTAL:** Reddito totale
- **AMT_CREDIT:** Importo del prestito
- **NAME_INCOME_TYPE:** Tipo di reddito
- **NAME_EDUCATION_TYPE:** Livello di educazione
- **NAME_FAMILY_STATUS:** Stato civile
- **NAME_HOUSING_TYPE:** Tipo di abitazione
- **OCCUPATION_TYPE:** Occupazione

Gruppo di Lavoro – Logistica

Coding

Siete gli analisti del reparto ecommerce di una grande catena di elettrodomestici in Italia che rivende sia piccoli che grandi elettrodomestici. L'ecommerce manager ha notato che l'algoritmo che calcola le spese di spedizione non funziona correttamente poichè segnava come costi di spedizione a tutti gli ordini 4,90€. Infatti i prodotti delle diverse categorie dell'ecommerce hanno diversi prezzi di spedizione e diverse scontistiche in base al prezzo totale. Ecco il listino:

- Piccoli elettrodomestici: Spese di spedizione 9,90€ - gratuite se il carrello totale supera 199€
- Grandi elettrodomestici: Spese di spedizione 29,90€ - gratuite se il carrello totale supera 499€
- Per i grandi elettrodomestici i clienti possono chiedere l'installazione del prodotto presso la propria abitazione al costo di 49,90€.

Create un algoritmo che calcoli correttamente le spese di spedizione, il costo totale e la differenza col prezzo attualmente applicato a partire da tre variabili:

- `tipologia_elettrodomestici` - stringa - può essere "piccolo" o "grande"
- `carrello_totale` - numero - valore in euro del carrello
- `installazione` - numero - 1 se il cliente richiede l'installazione 0 altrimenti

Big data

Nel settore dell'e-commerce, la puntualità delle consegne è cruciale per garantire la soddisfazione del cliente e mantenere un vantaggio competitivo. L'analisi dei dati consente di identificare i fattori che influenzano i ritardi nelle spedizioni, permettendo all'azienda di implementare strategie mirate per migliorare l'efficienza logistica. Ad esempio, molte aziende utilizzano modelli predittivi per anticipare possibili ritardi e ottimizzare le risorse in base alle esigenze operative.

In qualità di analisti nel reparto Business Intelligence di un'azienda internazionale di e-commerce specializzata nella vendita di prodotti elettronici, il vostro compito è esaminare i dati relativi alle spedizioni per identificare le variabili che incidono maggiormente sui ritardi nelle consegne. Attraverso questa analisi, potrete sviluppare modelli predittivi che consentano di anticipare i ritardi e implementare misure preventive. Questo approccio proattivo è essenziale per mantenere la soddisfazione del cliente e ottimizzare le operazioni logistiche. L'analisi dei dati vi permetterà di individuare pattern e tendenze che influenzano la puntualità delle consegne, fornendo all'azienda le informazioni necessarie per adottare strategie efficaci.

Obiettivo:

Identificare i fattori chiave che influenzano il ritardo delle consegne, al fine di implementare strategie efficaci per migliorare l'efficienza logistica e la soddisfazione del cliente.



Informazioni disponibili:

- **ID:** Identificativo univoco del cliente
- **Mode_of_Shipment:** Modalità di spedizione (Nave, Aereo, Strada)
- **Customer_rating:** Valutazione del cliente (1 = peggiore, 5 = migliore)
- **Cost_of_the_Product:** Costo del prodotto in dollari statunitensi
- **Prior_purchases:** Numero di acquisti precedenti
- **Product_importance:** Importanza del prodotto (bassa, media, alta)
- **Gender:** Genere del cliente (Maschio, Femmina)
- **Discount_offered:** Sconto offerto sul prodotto specifico
- **Weight_in_gms:** Peso del prodotto in grammi
- **Reached.on.Time_Y.N:** Variabile target (1 = consegna in ritardo, 0 = consegna puntuale)

Gruppo di Lavoro – Marketing

Coding

Siete il reparto data analyst di un noto franchising di moda low cost. Il responsabile marketing vuole lanciare una nuova offerta per il prossimo black friday. Per i clienti che acquistano almeno tre vestiti della categoria sport, il meno caro lo pagano solo 1€. Inoltre se la somma totale è superiore a 59€ lo scontrino viene scontato di un 5% aggiuntivo mentre se il totale supera 99€ lo sconto è pari al 10%.

Create un algoritmo che calcoli correttamente il prezzo totale a partire da un vettore con i prezzi dei tre prodotti.

Big Data

Analisi dei Dati per l'Ottimizzazione delle Campagne Promozionali

Nel settore della grande distribuzione, comprendere il comportamento dei clienti è fondamentale per sviluppare strategie di marketing efficaci. L'analisi dei dati consente di identificare i segmenti di clientela più propensi a rispondere positivamente alle offerte promozionali, ottimizzando così le risorse e massimizzando il ritorno sugli investimenti. Ad esempio, molte aziende utilizzano modelli predittivi basati su dati demografici e comportamentali per personalizzare le campagne e aumentare il tasso di risposta.

In qualità di analisti nel reparto Business Intelligence di un'azienda leader nella grande distribuzione, il vostro compito è esaminare i dati dei clienti per identificare i fattori che influenzano la probabilità di risposta alle offerte promozionali. Attraverso questa analisi, potrete sviluppare modelli predittivi che consentano di anticipare le reazioni dei clienti e implementare strategie di marketing mirate. Questo approccio proattivo è essenziale per migliorare l'efficacia delle campagne promozionali e aumentare la soddisfazione del cliente. L'analisi dei dati vi permetterà di individuare pattern e tendenze che influenzano le decisioni d'acquisto, fornendo all'azienda le informazioni necessarie per adottare strategie efficaci.

Obiettivo:

Identificare i fattori chiave che influenzano la probabilità di risposta positiva alle offerte promozionali, al fine di implementare strategie efficaci per migliorare l'efficacia delle campagne di marketing.

Informazioni disponibili:

- **Year_Birth**: Anno di nascita del cliente
- **Marital_Status**: Stato civile del cliente
- **Income**: Reddito annuale del cliente



- **Kidhome:** Numero di bambini piccoli nel nucleo familiare
- **Teenhome:** Numero di adolescenti nel nucleo familiare
- **MntWines:** Spesa per vini negli ultimi due anni
- **MntFruits:** Spesa per frutta negli ultimi due anni
- **MntMeatProducts:** Spesa per prodotti a base di carne negli ultimi due anni
- **MntFishProducts:** Spesa per prodotti ittici negli ultimi due anni
- **MntSweetProducts:** Spesa per dolci negli ultimi due anni
- **MntGoldProds:** Spesa per prodotti in oro negli ultimi due anni
- **Response:** 1 se il cliente ha accettato l'offerta nell'ultima campagna, 0 altrimenti (variabile target)



Gruppo di Lavoro – Sanità

Coding

Siete i data analyst di una ONG che opera in Africa. Il responsabile del reparto medico vi chiede di creare un'app da distribuire a tutti gli ospedali da campo per velocizzare la diagnosi della febbre tropicale dengue. Vi segnala una ricerca scientifica (<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0161696>) in cui sono stati pubblicati i parametri per classificare il rischio elevato o ridotto del paziente in base al sintomo di vomito, versamento pleurico e la pressione sanguigna sistolica:

- Se il paziente ha sintomi di vomito il rischio è elevato
- Se il paziente ha versamenti pleurici il rischio è elevato
- Se il paziente ha pressione sanguigna sistolica inferiore a 90 (mmHg) il rischio è elevato
- In tutti gli altri casi il rischio è ridotto

Create un algoritmo che restituisca correttamente la classificazione di rischio del paziente.

Big Data

Siete nel reparto Business Intelligence di una azienda sanitaria. Il dataset contiene i dati di alcuni pazienti che hanno riferito di avere una malattia coronarica o un infarto del miocardio. Quali informazioni potrà estrarre l'azienda per valutare il rischio di malattie cardiache?

Le informazioni a disposizione sono:

- HeartDisease - Intervistati che hanno mai riferito di avere una malattia coronarica (CHD) o un infarto del miocardio (IM)
- BMI - BMI
- Smoking - Fuma (si/no)
- AlcoholDrinking - Beve alcool (si/no)
- Stroke - Ictus
- PhysicalHealth - Salute fisica, che include malattie fisiche e lesioni, per quanti giorni negli ultimi 30 giorni la salute fisica non è stata buona? (0-30 giorni)
- MentalHealth - Salute mentale, per quanti giorni negli ultimi 30 giorni la salute mentale non è stata buona? (0-30 giorni)
- DiffWalking - Difficoltà nel camminare
- Sex - Sesso
- AgeCategory - Fascia età
- Race - Razza
- Diabetic - Diabete
- PhysicalActivity - Attività fisica
- GenHealt - Opinione della salute in generale



- SleepTime - Ore di sonno giornaliera
- Asthma - Asma
- KidneyDisease - Malattie renali
- SkinCancer - Cancro della pelle