

Previsione dell'affidabilità creditizia per il rilascio della carta di credito

Sei stato assunto dalla Pro National Bank come data scientist, il tuo primo incarico consiste nel realizzare un modello in grado di stimare l'affidabilità creditizia di un clienti, al fine di di aiutare il team dedicato a comprendere se accettare o meno la richiesta per il rilascio della carta di credito.

A tal fine ti vengono consegnati i dati anonimizzati di clienti che hanno già ottenuto la carta di credito e lo stato del loro debito mensile.

I dati sono organizzati in 2 file csv che possono essere scaricati da qui.

Il file [application_type.csv](#) contiene le informazioni dei correntisti che hanno richiesto l'apertura di una linea di credito.

- **ID:** numero identificativo del cliente
- **CODE_GENDER:** sesso del cleinte
- **FLAG_OWN_CAR:** indicatore del possesso di un'automobile
- **FLAG_OWN_REALTY:** indicatore del possesso di una casa
- **CNT_CHILDREN:** numero di figli
- **AMT_INCOME_TOTAL:** reddito annuale
- **NAME_INCOME_TYPE:** tipo di reddito
- **NAME_EDUCATION_TYPE:** livello di educazione
- **NAME_FAMILY_STATUS:** Stato civile

- **NAME_HOUSING_TYPE:**
- **DAYS_BIRTH:** Numero di giorni trascorsi dalla nascita
- **DAYS_EMPLOYED:** Numero di giorni trascorsi dalla data di assunzione, se positivo indica il numero di giorni da quando è disoccupato
- **FLAG_MOBIL:** indicatore della presenza di un numero di cellulare
- **FLAG_WORK_PHONE:** indicatore della presenza di un numero di telefono di lavoro
- **FLAG_PHONE:** indicatore della presenza di un numero di telefono
- **FLAG_EMAIL:** indicatore della presenza di un indirizzo email
- **OCCUPATION_TYPE:** tipo di occupazione
- **CNT_FAM_MEMBERS:** numero di familiari

Il file credit_record.csv contiene lo storico dei correntisti che hanno una linea di credito attiva.

- **ID:** numero identificativo del cliente
- **MONTHS_BALANCE:** Il mese dei dati estratti rappresenta il punto di partenza, a ritroso: 0 è il mese corrente, -1 è il mese precedente, e così via.
- **STATUS:** Valore 0: da 1 a 29 giorni di ritardo. Valore 1: da 30 a 59 giorni di ritardo. Valore 2: da 60 a 89 giorni di ritardo. Valore 3: da 90 a 119 giorni di ritardo. Valore 4: da 120 a 149 giorni di ritardo. Valore 5: Debiti scaduti o insolventi, cancellazioni per più di 150 giorni. Valore C: saldato nel mese. Valore X: Nessun prestito per il mese

Il tuo modello deve essere in grado di fornire un'indicazione al team nell'approvare o meno la richiesta di carta di credito per un dato cliente.

PUNTO BONUS

Se ad un cliente viene negata la carta di credito, il team deve essere in grado di fornirgli una motivazione, questo vuol dire che il tuo modello deve fornire delle indicazioni facilmente interpretabili.

NOTE

Come avrai notato manca una variabile target, non ti vogliamo dare indicazioni, sta a te capire come procedere. Qualora non riuscissi a venirne a capo chiedi pure una mano al tuo coach nella classe virtuale di Machine Learning su Discord.