

Coding e Big Data

2024-2025



Vincenzo Nardelli

vincenzo.nardelli@unicatt.it

Questo dataset è stato raccolto casualmente dal database di una compagnia telefonica in un periodo di 12 mesi.

In totale, sono presenti 3150 righe di dati, ognuna delle quali rappresenta un cliente e contiene informazioni per 13 colonne

Input:

- **Fallimenti delle Chiamate:** numero di chiamate fallite
- **Reclami:** variabile binaria (0: nessun reclamo, 1: reclamo)
- **Durata dell'Abbonamento:** totale dei mesi di abbonamento
- **Importo Addebitato:** attributo ordinale (0: importo più basso, 9: importo più alto)
- **Secondi di Utilizzo:** totale dei secondi di chiamate
- **Frequenza d'Uso:** numero totale di chiamate
- **Frequenza di SMS:** numero totale di messaggi di testo inviati
- **Numero di Chiamate Distinte:** numero totale di chiamate effettuate a numeri distinti
- **Gruppo di Età:** attributo ordinale (1: età giovane, 5: età avanzata)
- **Piano Tariffario:** variabile binaria (1: tariffa a consumo, 2: contratto)
- **Stato:** variabile binaria (1: attivo, 2: non attivo)
- **Churn:** variabile binaria (1: abbandono, 0: non abbandono) - Etichetta di classe
- **Valore del Cliente:** valore calcolato del cliente