


GABRIEL BARDAJÍ BIESCAS 1568297

HEALTH INSURANCE CROSS SELL PREDICTION

APRENTATGE COMPUTACIONAL

- 
- 01 INTRODUCCIÓ AL DATASET**
 - 02 ANÀLISI DELS ATRIBUTS**
 - 03 APLICACIÓ DELS MÈTODES D'APRENENTATGE**
 - 04 VISUALITZACIÓ DELS MÈTODES D'APRENENTATGE**
 - 05 CONCLUSIONS**

INTRODUCCIÓ AL DATASET

- Recopilació de dades sobre els clients d'una asseguradora de vehicles.
- **381.109** files i **12** columnas

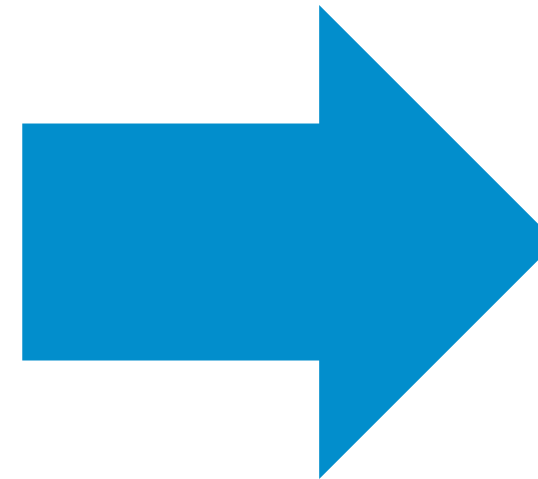


| Gender | Age | Driving_License | Region_Code | Previously_Insured | Vehicle_Age | Vehicle_Damage | Annual_Premium | Policy_Sales_Channel | Vintage | Response |
|--------|-----|-----------------|-------------|--------------------|-------------|----------------|----------------|----------------------|---------|----------|
| Male | 44 | 1 | 28.0 | 0 | > 2 Years | Yes | 40454.0 | 26.0 | 217 | 1 |
| Male | 76 | 1 | 3.0 | 0 | 1-2 Year | No | 33536.0 | 26.0 | 183 | 0 |
| Male | 47 | 1 | 28.0 | 0 | > 2 Years | Yes | 38294.0 | 26.0 | 27 | 1 |
| Male | 21 | 1 | 11.0 | 1 | < 1 Year | No | 28619.0 | 152.0 | 203 | 0 |
| Female | 29 | 1 | 41.0 | 1 | < 1 Year | No | 27496.0 | 152.0 | 39 | 0 |



ANÀlisi DELS ATRIBUTS

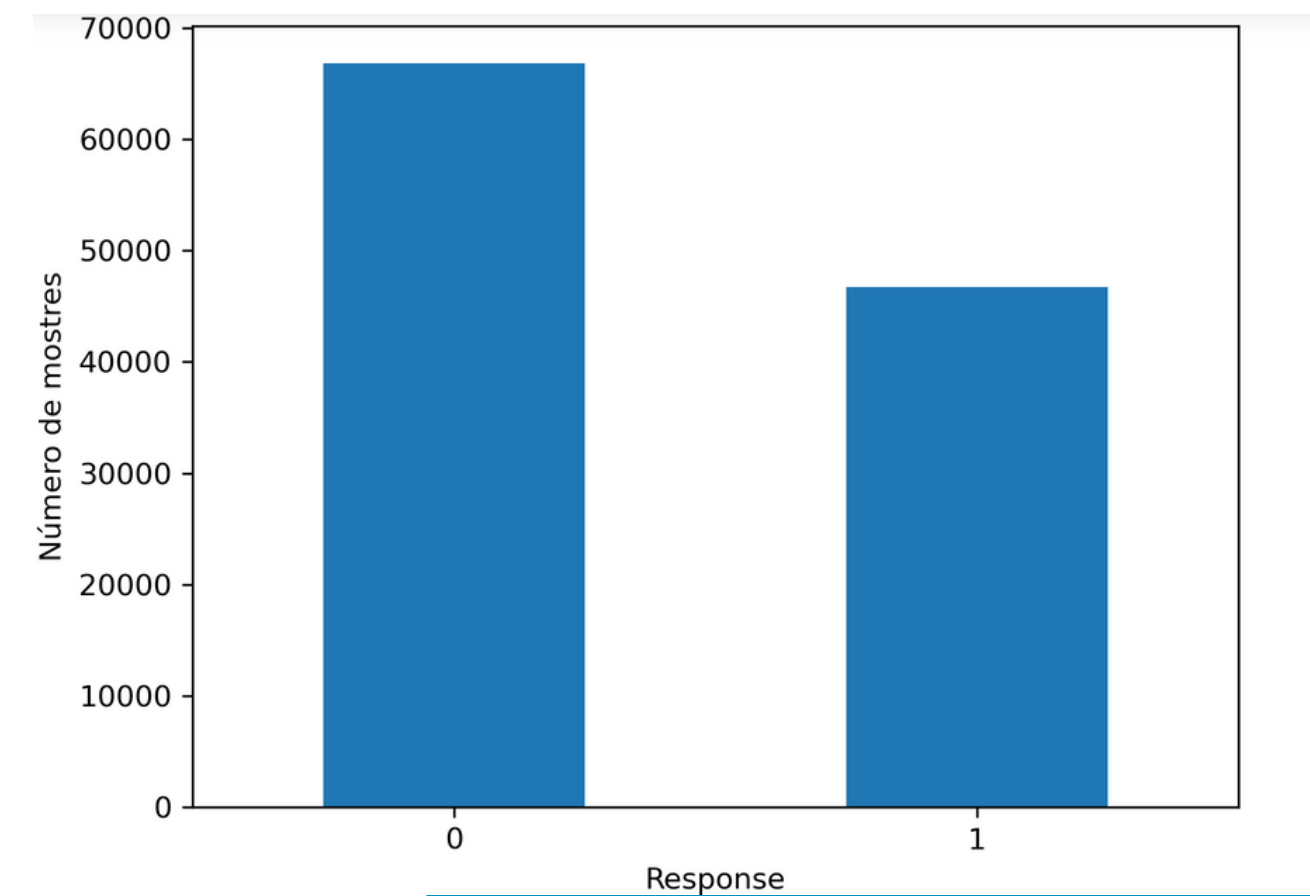
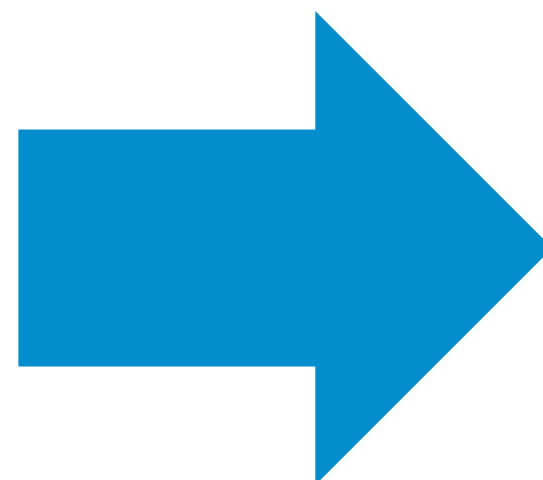
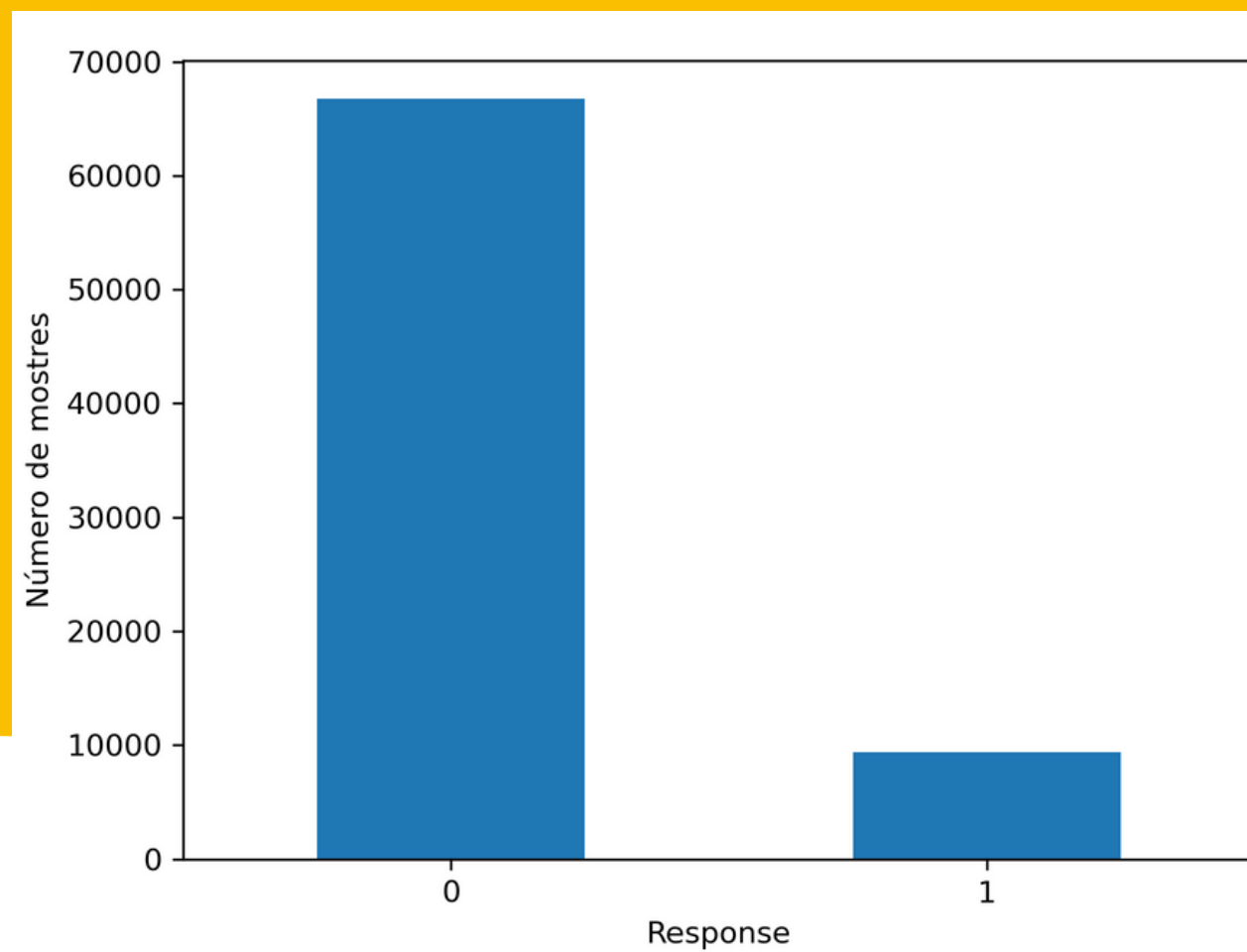
| | |
|----------------------|---------|
| id | int64 |
| gender | object |
| age | int64 |
| driving_license | int64 |
| region_code | float64 |
| previously_insured | int64 |
| vehicle_age | object |
| vehicle_damage | object |
| annual_premium | float64 |
| policy_sales_channel | float64 |
| vintage | int64 |
| response | int64 |



| | |
|----------------------|---|
| id | 0 |
| Gender | 0 |
| Age | 0 |
| Driving_License | 0 |
| Region_Code | 0 |
| Previously_Insured | 0 |
| Vehicle_Age | 0 |
| Vehicle_Damage | 0 |
| Annual_Premium | 0 |
| Policy_Sales_Channel | 0 |
| Vintage | 0 |
| Response | 0 |
| dtype: int64 | |
| (381109, 12) | |



ANÀLISI DELS ATRIBUTS





ANÀlisi DELS ATRIBUTS - TRACTAMENT DE VARIABLES CATEGÒRIQUES

1

VEHICLE_AGE

- '> 2 Years' = 0
- '1-2 Year' = 1
- '< 1 Year' = 2

2

VEHICLE_DAMAGE

- Yes = 1
- No = 0

3

GENDER

- Male = 1
- Female = 0

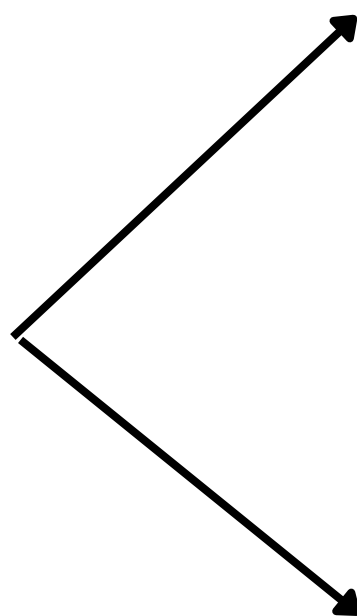


MODEL SELECTION

1 DESBALANCEJAT

2 BALANCEJAT

3 SMALL-DATASET



TOTS ELS ATRIBUTS

**ATRIBUTS AMB MÉS
CORRELACIÓ**



VISUALITZACIÓ DELS MÈTODES D'APRENENTATGE

| | DESBALANCEJADA | | BALANCEJADA | | SMALL_DATASET | |
|-------------------|-------------------|----------|-------------------|----------|-------------------|----------|
| | TOTS ELS ATRIBUTS | MÉS CORR | TOTS ELS ATRIBUTS | MÉS CORR | TOTS ELS ATRIBUTS | MÉS CORR |
| KNN | 0.8752 | 0.8760 | 0.7685 | 0.7721 | 0.8750 | 0.8673 |
| Regresor Logístic | 0.8770 | 0.8768 | 0.7449 | 0.7449 | 0.8728 | 0.8730 |
| Gradient Descent | 0.8774 | 0.8775 | 0.7599 | 0.7584 | 0.8743 | 0.8738 |
| Random Forest | 0.8655 | 0.7253 | 0.7614 | 0.7692 | 0.8655 | 0.7748 |
| Naive Bayes | 0.63833 | 0.6461 | 0.7490 | 0.3380 | ----- | ----- |



CONCLUSIONS



ENLLAÇ AL GITHUB



https://github.com/biel1703/Kaggle_individual_ApC

**GRÀCIES PER LA
VOSTRA ATENCIÓ**