# CREDIT CARD FRAUD DETECTION

**Autor:** Manel Carrillo Maíllo

## 0.Abstract

Aquest estudi es centra a abordar el desafiament del desbalanceig de dades de transaccions financeres mitjançant l'ús de tècniques d'oversampling i undersampling. A més, s'explora la utilització d'autoencoders amb l'objectiu de detectar outliers amb una major precisió.

## 1. Descripció del conjunt de dades

En aquesta investigació s'ha utilitzat el dataset https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud que consta de 284.807 transaccions amb un total de 31 columnes.

Aquest dataset representa les transaccions fetes amb targeta de crèdit durant el setembre de 2013 a Europa, en un període de dos dies.

Les columnes principals del conjunt de dades inclouen:

**Time:** Temps transcorregut entre la primera transacció i la transacció actual.

**Amount**: Indica la quantitat monetària associada amb cada transacció.

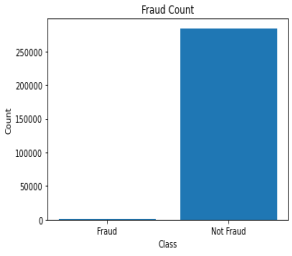
**Class:** Variable binària que classifica les transaccions com a "No frau" (0) o "Frau" (1).

**V [1,29]):** Són les components principals obtingudes d' un anàlisi de components principals (PCA). La informació específica d'aquestes components es manté confidencial per raons de privadesa.

## 2. Preprocessament i Problema a Resoldre

El conjunt de dades és inicialment senzill pel que fa al preprocessament, ja que no haurem de fer cap mena d'encoding pel fet que no conté valors nuls ni cap variable categòrica.

No obstant això, cal destacar un aspecte crític en aquesta secció, que esdevé el principal problema a tractar: el desbalanceig elevat de classes.




Només un 0,17% (492) de transaccions estan classificades com a frau mentre que les transaccions no fraudulentes, representen un 99,83% (284.315) de les dades.

Solucionar de manera efectiva aquest desequilibri és fonamental per garantir la capacitat de detecció de transaccions fraudulentes dels futurs models.

En aquest estudi, per buscar aquest equilibri, s'han implementat tècniques d'oversampling com la tècnica SMOTE (Synthetic Minority Over-sampling Technique) i ROS (Random Over Sampling), així com tècniques d'undersampling com RUS (Random Under-Sampling) o NearMiss.

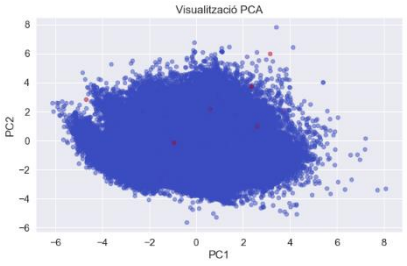
Aquest desbalanceig tan pronunciat serà el principal repte a superar a l'hora de construir models de detecció de frau.

Solucionar de manera efectiva aquest desequilibri és fonamental per garantir la capacitat de detecció de transaccions fraudulentes dels models.

En aquest estudi, per contrarestar aquest equilibri, s'han implementat amb èxit tècniques d'oversampling com la tècnica SMOTE (Synthetic Minority Over-sampling Technique) i ROS (Random Over Sampling), així com tècniques d'undersampling com RUS (Random Under-Sampling) o NearMiss.

## 3. Aplicació de PCA i Emplenament amb Tècniques d'Oversampling i Undersampling

Amb l'objectiu de comprendre millor la distribució de dades del dataset desbalancejat, es va dur a terme una anàlisi de components principals o PCA.




Com es va mencionar a l'apartat anterior, aquest PCA mostra de manera visual la gran desproporció entre les classes del dataset.

Posteriorment, per gestionar eficaçment la desigualtat en la distribució de classes, es van aplicar tècniques d'oversampling i undersampling amb l'objectiu de trobar la tècnica que millori la precisió de detecció dels futurs models i genereat els seus respectius PCA per veure com funcionen a la pràctica.

### 3.1 Tècniques D'oversampling

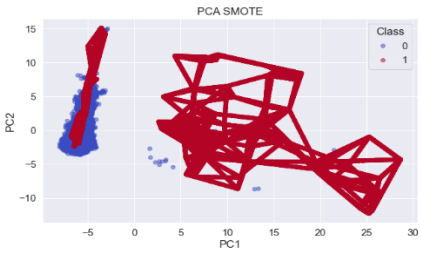
Les tècniques d'oversampling es basen en generar noves dades de la classe minoritària amb l'objectiu d'equilibrar les classes del dataeset.

**SMOTE (Synthetic Minority Over-sampling Technique)**

Aquesta tècnica genera mostres sintètiques de la classe minoritària mitjançant un procés d'interpolació entre els punts ja existents d'aquesta classe.

La tècnica SMOTE, selecciona una instància de la classe minoritària i identifica els seus veïns més pròxims , també de classe minoritària (KNN).

A continuació, s'introdueixen noves instàncies sintètiques entre l'observació original i aquests veïns, creant una major densitat de dades.




**ROS (Random Over-Sampling):**

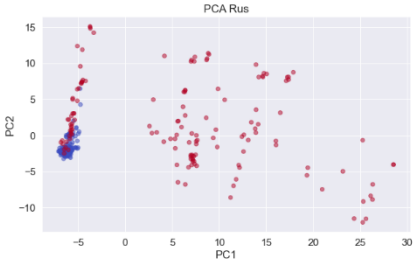
En paral·lel amb l'ús de SMOTE, s'ha aplicat Random Over-Sampling (ROS) per gestionar el desequilibri entre les classes. En aquesta tècnica, es dupliquen les mostres de la classe minoritaria a partir d'una selecció aleatòria de mostres d'aquesta classe.


### 3.2 Tècniques d'Undersampling

En contrast de les tècniques d'oversampling, s'ha dut a terme tècniques d'undersampling per gestionar l'abundància de la classe majoritària eliminant instàncies d'aquesta.

**RUS (Random Under-sampling)**

Aquesta tècnica, selecciona aleatòriament mostres de la classe majoritària i s'eliminen per reduir la quantitat de transaccions d'aquesta classe buscant l'equilibri a la distribució de les classes.

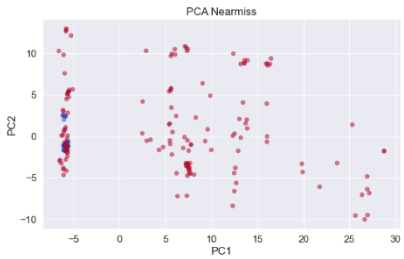



**Near Miss**

En aquest mateix context de buscar l'equilibri de classes a partir de l'eliminació d'instàncies de classe majoritària, s'ha aplicat la tècnica de NearMiss.

Aquesta tècnica selecciona mostres de la classe majoritària basant-se en la seva proximitat amb les instàncies de la classe minoritària (KNN) i elimina les mostres similars de la classe majoritària que podrien no afegir informació al model.

És a dir, elimina aquelles transaccions classificades com 0 que són més properes a transaccions de la classe 1




## 4. Estudi de Selecció de Mètriques

En un problema de classificació, com és el cas en la detecció de frau amb targetes de crèdit, i més quan les classes estan desbalancejades, la tria adequada de mètriques és molt important per a una avaluació precisa del rendiment del model.

En aquest estudi, hem reconegut les limitacions de l'accuracy a datasets desbalancejats, ja que pot predir molt bé les dades de la classe majoritària però fallar a la classe minoritària i igualment donar valors molt alts degut al desbalanceig.

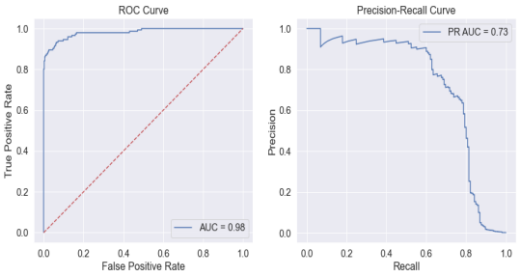
Per tant, hem optat per centrar-nos en mètriques més informatives, com el f1-score i la corba PR (Precision-Recall).

Aquestes mètriques proporcionen una comprensió més gran de la capacitat del model en la detecció de la classe minoritària en situacions de classes desbalancejades.

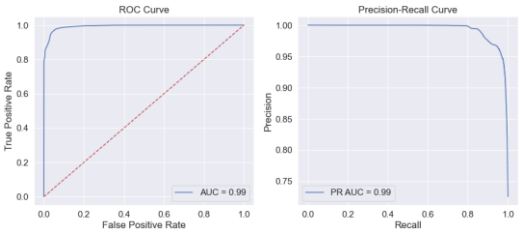
Per a fer una primera avaluació de les tècniques d'equilibri de dades i per valorar de manera general les mètriques, hem generat un logistic regression model senzill i generant les corbes ROC i PR.

Aquest regression model s'ha entrenat amb el dataset original amb el conjunt d'entrenament i posteriorment amb els datasets posteriors a aplicar les tècniques d'equilibri de classes amb la finalitat de poder fer una primera comparació.

Si ens fixem a les corbes ROC i PR del dataset desbalancejat, podem veure que té un AUC a la ROC de gairebé 1 però un AUC de 0.73 a la corba PR indicant que té dificultats per identificar la classe positiva.




Per altra banda, tots els models, tant els d'oversampling com undersampling tenen un AUC de 0.99 tant a la corba ROC com a la corba PR indicant que poden detectar a la perfecció els fraus.




## 5. Validació Creuada dels Models

Per aprofundir en l'avaluació dels models, s'ha realitzat una validació creuada estratificada amb 5 parts, utilitzant les dades d'entrenament modificades amb cada tècnica d'oversampling i undersampling amb l'objectiu de comparar diversos models de classificació per determinar el seu rendiment en termes de la puntuació F1.

### 5.1 Models Avaluats

Els següents models han estat sotmesos a la validació creuada:

Logistic Regression, Random Forest, Naive Bayes, AdaBoost, Gradient Boosting.

### 5.2 Resultats Validació creuada al Train set

Els millors resultats que hem obtingut han estat als mètodes d'oversampling (SMOTE i ROS)

**Smote** **ROS**

| Model | F1 | Model | F1 |
|---|---|---|---|
| R.Forest | 0.999 | R.Forest | 0.999 |
| Grad.Boosting | 0.993 | Grad.Boosting | 0.998 |

## 6. Avaluacio del Millor Model al Test sample

Gràcies a la informació extreta dels resultats obtinguts a partir de la validació creuada, sabem que el millor model per a la tècnica SMOTE i ROS, és el random forest.

Per tal d'avaluar la generalització del model, s'ha provat amb les dades del test generant el classification report i més tard les corbes PR i ROC, de les quals ens interessa més avaluar la corba PR ja que relaciona la precision i la recall, mètriques presents al càlcul del f1-score.

A continuació, es mostren els resultats clau d'aquesta avaluació :

Per a la tècnica SMOTE:

RandomForestClassifier - Classification Report:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 1.00 | 1.00 | 1.00 | 39289 |
| 1 | 1.00 | 0.78 | 0.88 | 36 |
| ... | | | | |
| accuracy | | | 1.00 | 39325 |
| macro avg | 1.00 | 0.89 | 0.94 | 39325 |
| weighted avg | 1.00 | 1.00 | 1.00 | 39325 |

Com podem veure, el model demostra una excepcional capacitat de detecció de fraus, amb una precision del 100% i un recall del 78% per a

la classe 1. L'F1-score del 88% a la classe minoritària indica un alt rendiment global per detectar fraus.

Per a la tècnica ROS:

```
RandomForestClassifier - Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     39289
           1       1.00      0.69      0.82        36
...
    accuracy                           1.00     39325
   macro avg       1.00      0.85      0.91     39325
weighted avg       1.00      1.00      1.00     39325
```

Podem dir que respecte la tècnica SMOTE, si analitzem la recall de la classe 1 i el f1-score, la tècnica ROS funciona lleugerament pitjor tot i que també dona resultats bons.

Als dos classification report, ens apareix que el weighted avg f1-score és de 1.

El wheighted avg f1-score dona és importància a les classes amb més instàncies vertaderes fet que és útil en cassos com aquest in hi ha desbalanceig de classes. De manera que obtenir un wheighted f1-score tant alt, és molt bona senyal ja que indica la bona capacitat de generalització del model.

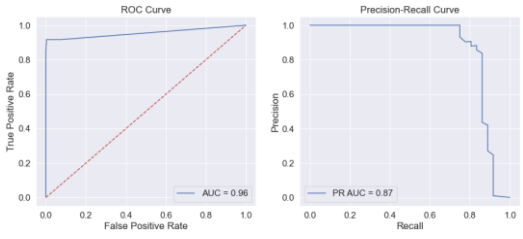
## 7 Anàlisis final i Reflexions

Per obtenir una comprensió més detallada del rendiment del model seleccionat en el conjunt de test, s'ha realitzat una anàlisi final que inclou diverses visualitzacions rellevants com les corbes PR i ROC, un gràfic de classes predites vs probabilitat de la classe 1 i la matriu de confusió.

Tot això només dels models que més bons resultats han donat en el nostre estudi, l'SMOTE i ROS.

### 7.1 Corbes ROC i PR

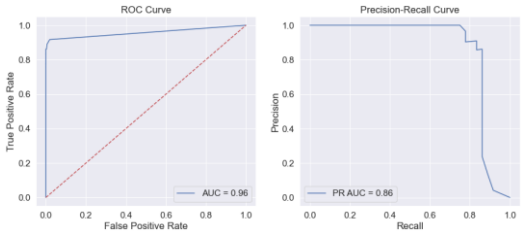
Per al model SMOTE tenim un AUC de la corba ROC de 96 i un AUC de la corba PR de 87.




A partir d'aquests resultats, el nostre model RandomForest entrenat amb les dades modificades amb la tècnica SMOTE, demostra un sòlid equilibri entre precisió (100%) i recall (78%) en la detecció de fraus.

Aquest alt rendiment indica una eficaç capacitat per identificar transaccions fraudulentes.

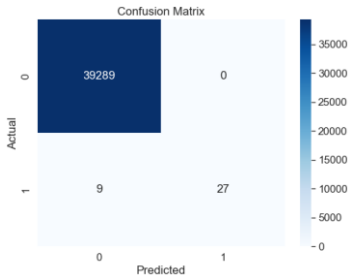
El mateix podem dir del model ROS que presenta un AUC de 86 a la corba PR.



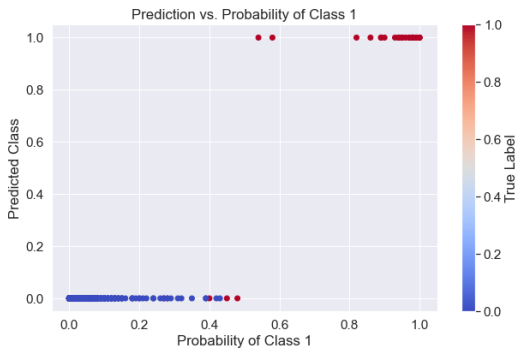

### 7.3 Matriu de Confusió i Grafic classes predites vs probabilitats

La matriu de confusió proporciona una visió més detallada de les prediccions del model en el conjunt de test. Aquesta taula mostra els veritables positius, veritables negatius, falsos positius i falsos negatius.

La matriu de confussió per al model SMOTE és:



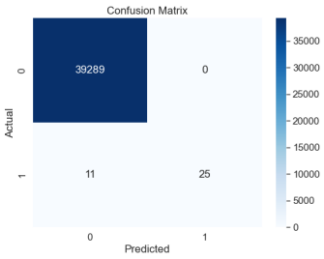

On podem veure que classifica bé totes les instàncies de classe 0 i les de classe 1 classifica bé 27 de 36 demostrant una gran capacitat a l'hora de classificar i detectar fraus a les transaccions.



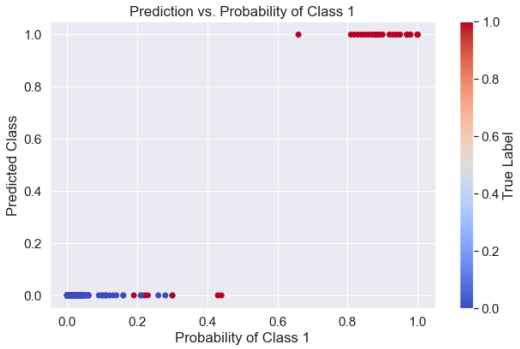

Podem veure de manera visual la mateixa informació que hem mencionat abans on es veu

que hi ha petits errors a la predicció de la classe 1.

Per altra banda, la matriu de confussió del model ROS és:




On al igual que el model SMOTE, classifica bé les instàncies de classe 0 però de les de classe 1, en classifica bé 25 de 36, sent lleugerament superior el model SMOTE que el model ROS.



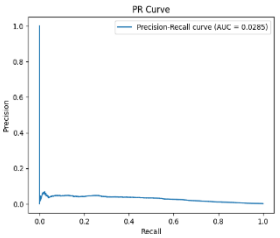

Si ens fixem i comparem aquest plot de prediccions contra la probabilitat de la classe 1 amb el de la tècnica SMOTE, podem veure que hi ha més instàncies de classe 1 mal predites.

## 8. Model Autoencoder per a la Detecció d'Anomalies

A més dels models de classificació convencionals, hem explorat l'ús d'un autoencoder com a eina per a la detecció d'anomalies en les transaccions amb targetes de crèdit. Aquest model, tot i no obtenir resultats finals òptims, ha ofert una perspectiva única sobre la naturalesa de les dades.

Hem avaluat el model autoencoder utilitzant el mean squared error (MSE) entre les dades originals i les reconstruccions del model.

A continuació, s'han utilitzat les puntuacions de MSE per generar un dataset juntsament amb les classes del dataset per poder construir les corbes ROC i PR




Com podem veure, tenim un AUC a la corba PR de 0.0285, això vol dir que el model autoencoder no ha estat efectiu en la identificació d'anomalies i no és capaç de distingir les classes.

Així, malgrat no ser la solució definitiva, l'ús de l'autoencoder ha proporcionat una exploració addicional que pot ser prometedora amb ajustaments i refinaments futurs.

## 9 Conclusió Final

A través d'aquesta investigació, hem abordat amb èxit el desafiant problema de la detecció de fraus en transaccions amb targetes de crèdit.

Mitjançant l'ús d'un model RandomForest entrenat amb el dataset modificat per la tècnica SMOTE, hem superat el desbalanceig de dades, aconseguint un equilibri notable entre precisió (100%) i recall (78%) amb un f1-score alt.

Aquesta eficàcia demostra la solidesa del model per identificar amb precisió les transaccions fraudulentes, a l'hora que minimitza els falsos positius.

Per altra banda, podem finalitzar dient, que en aquest estudi, els models entrenats amb datasets equilibrats amb tècniques d'oversampling han estat molt superiors als equilibrats amb tècniques d'undersampling, segurament per la pèrdua d'informació del dataset original al reduir tant les dimensions de les dades.