

# Mecanismo para Detecção de Transações de Cartão de Crédito Fraudulentas

Douglas Affonso Clementino  
Rafael De Paulo Dias





# Objetivo

Efetuar a classificação de transações de cartão de crédito com foco na detecção de **Transações Fraudulentas** (recall).



# Base de Dados

## Credit Card Fraud Detection

Anonymized credit card transactions  
labeled as fraudulent or genuine

**kaggle**



# Base de Dados

## Credit Card Fraud Detection

Anonymized credit card transactions  
labeled as fraudulent or genuine

**kaggle**

## Variáveis

- Time
- Amount
- Class
- v1-v28



# Base de Dados

## Credit Card Fraud Detection

Anonymized credit card transactions  
labeled as fraudulent or genuine

**kaggle**

## Variáveis

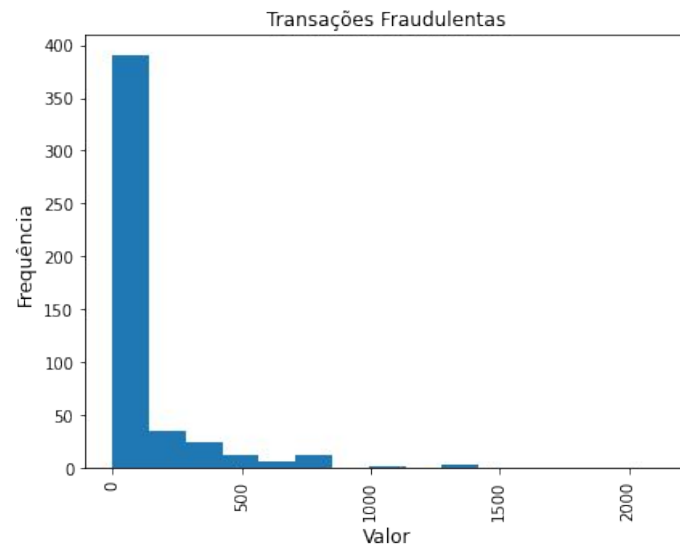
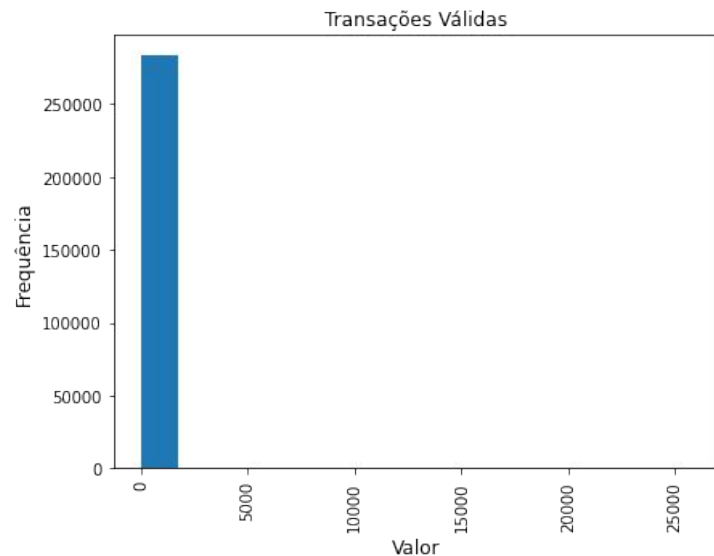
- Time
- Amount
- Class
- v1-v28

## Distribuição

- Total de Amostras: 284.807
- Classe 0 (Válidas): 284.315 (99,82%)
- Classe 1 (Fraude): 492 (0,17%)

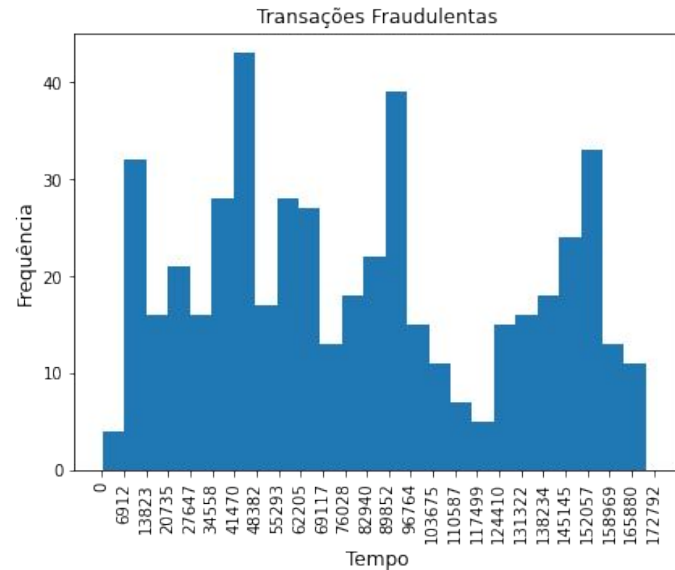
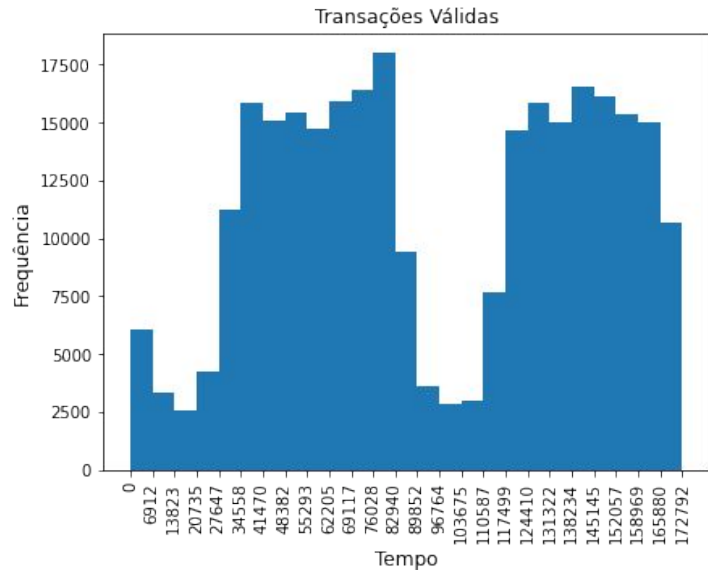


# Análise de Dataset





# Análise de Dataset





# Experimentos

## Escolha de Classificador

- KNeighborsClassifier (KNN)
- Multi-layer Perceptron (MLP)
- CatBoostClassifier (Random Forest)





# Experimentos

## Escolha de Classificador

- KNeighborsClassifier (KNN)
- Multi-layer Perceptron (MLP)
- CatBoostClassifier (Random Forest)

## Tuning de Classificador (CatBoostClassifier)

- **learning\_rate:** 0.0947
- **max\_depth:** 7
- **n\_estimators:** 700



# Experimentos

## Escolha de Classificador

- KNeighborsClassifier (KNN)
- Multi-layer Perceptron (MLP)
- CatBoostClassifier (Random Forest)

## Tuning de Classificador (CatBoostClassifier)

- `learning_rate`: 0.0947
- `max_depth`: 7
- `n_estimators`: 700.0

## Balanceamento de Bases de Treino

- Synthetic Minority Over-sampling (SMOTE)



# Experimentos

## Escolha de Classificador

- KNeighborsClassifier (KNN)
- Multi-layer Perceptron (MLP)
- CatBoostClassifier (Random Forest)

## Tuning de Classificador (CatBoostClassifier)

- `learning_rate`: 0.0947
- `max_depth`: 7
- `n_estimators`: 700.0

## Balanceamento de Bases de Treino

- Synthetic Minority Over-sampling (SMOTE)

## Considerações

- Kfold
- Outliers
- CatBoostClassifier GPU

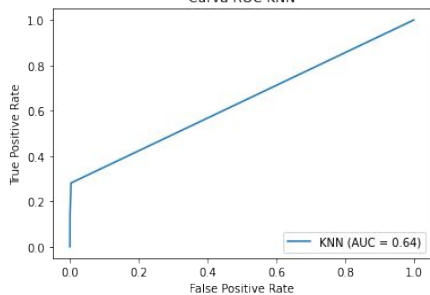


# **Escolha de Classificador**

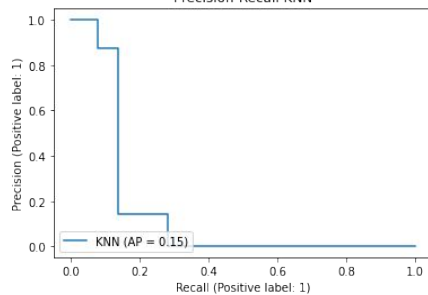


# Escolha de Classificador

Curva ROC KNN



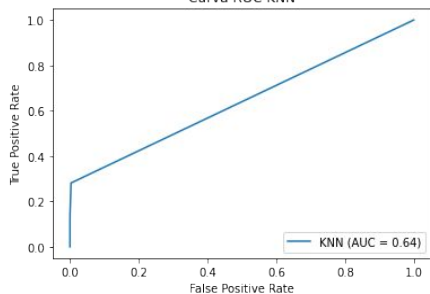
Precision-Recall KNN



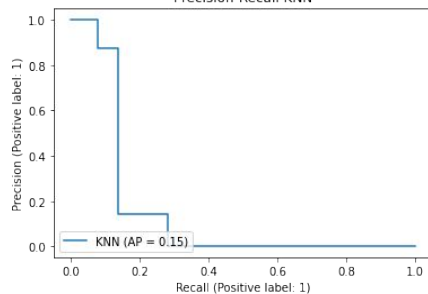


# Escolha de Classificador

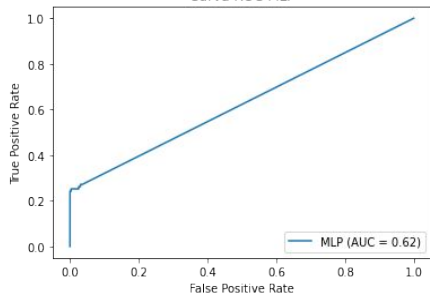
Curva ROC KNN



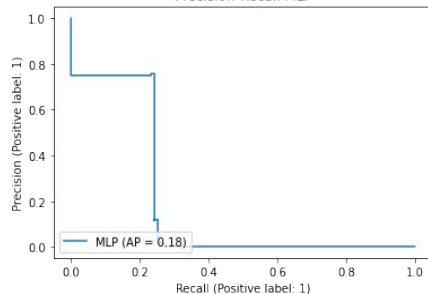
Precision-Recall KNN



Curva ROC MLP



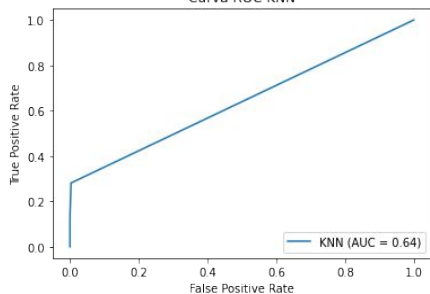
Precision-Recall MLP



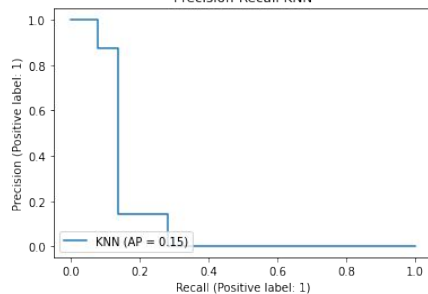


# Escolha de Classificador

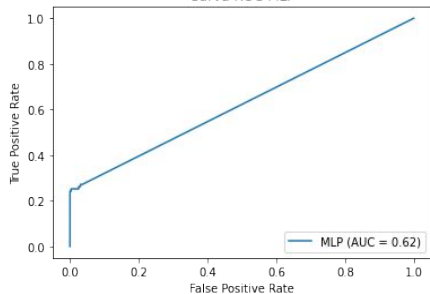
Curva ROC KNN



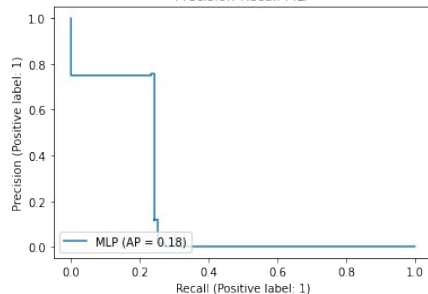
Precision-Recall KNN



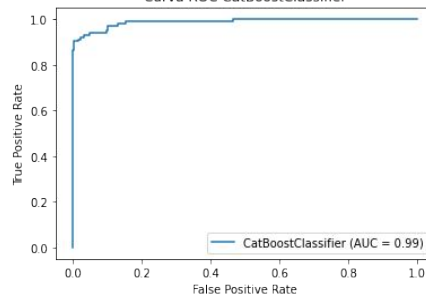
Curva ROC MLP



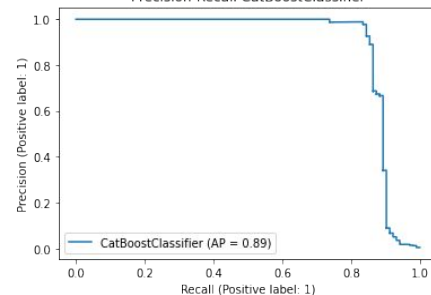
Precision-Recall MLP



Curva ROC CatBoostClassifier



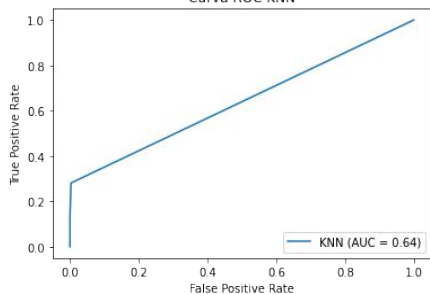
Precision-Recall CatBoostClassifier



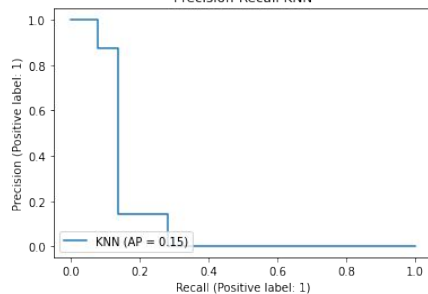


# Escolha de Classificador

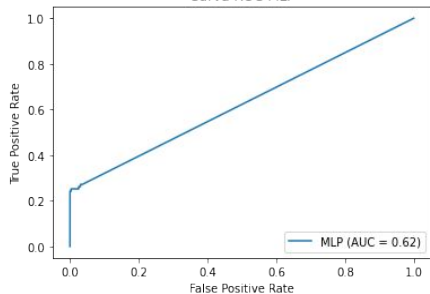
Curva ROC KNN



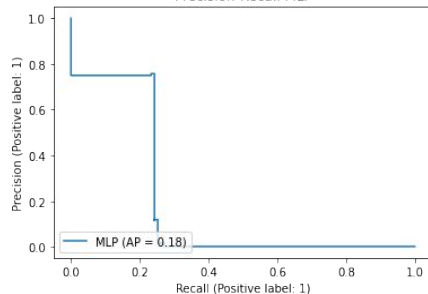
Precision-Recall KNN



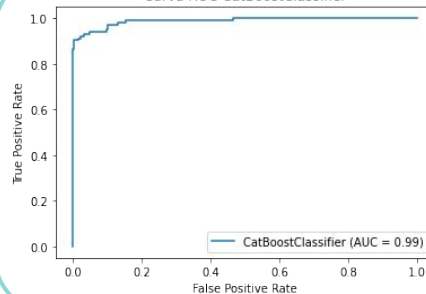
Curva ROC MLP



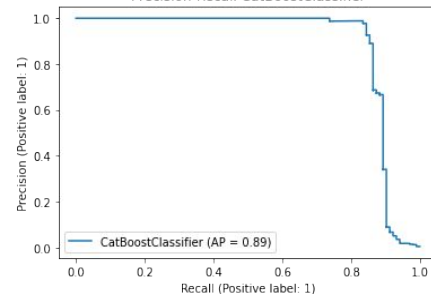
Precision-Recall MLP



Curva ROC CatBoostClassifier



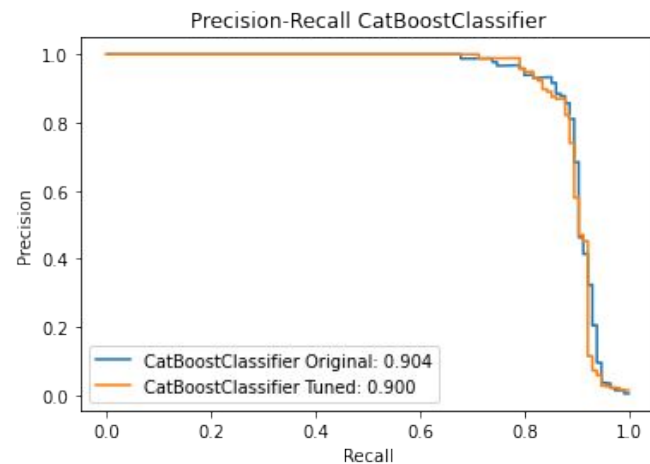
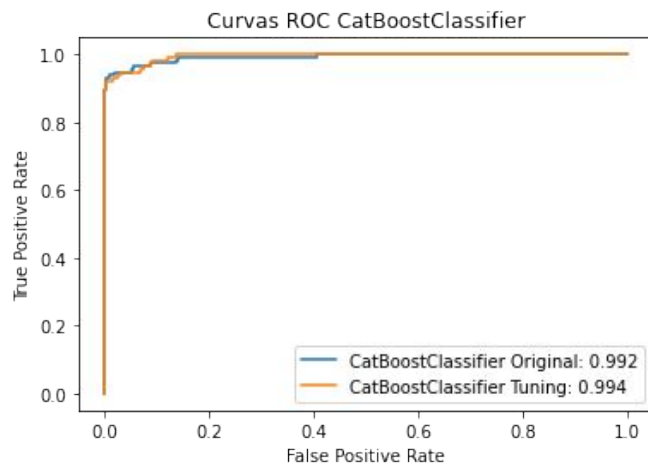
Precision-Recall CatBoostClassifier





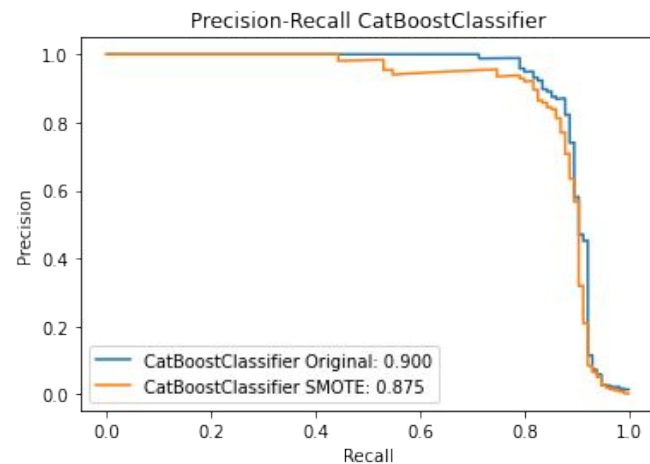
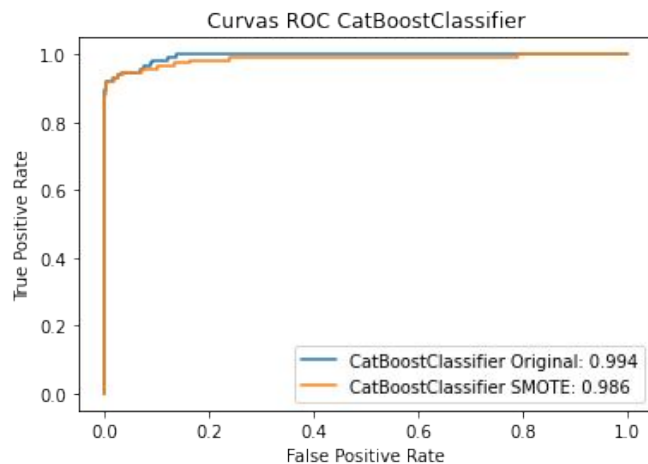


# Tuning de Classificador



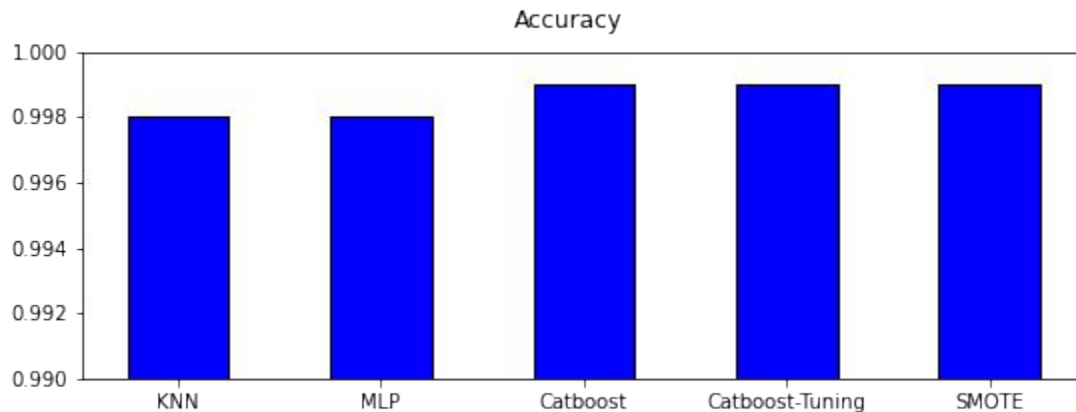


# Over-sampling (SMOTE)



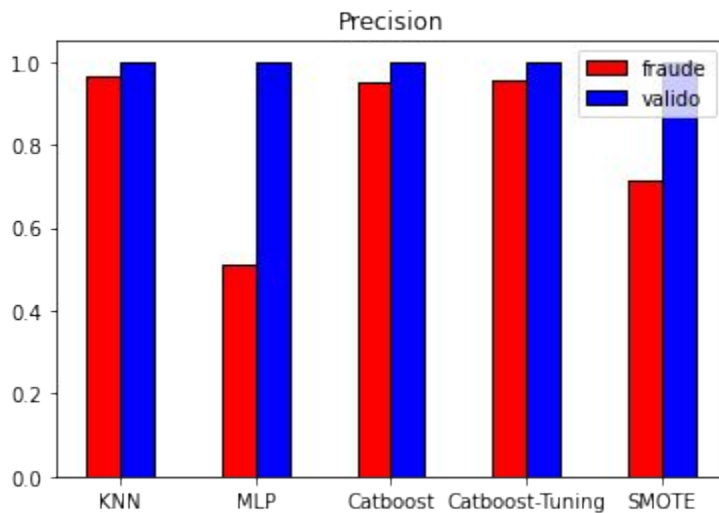


# Médias de Métricas por Modelo



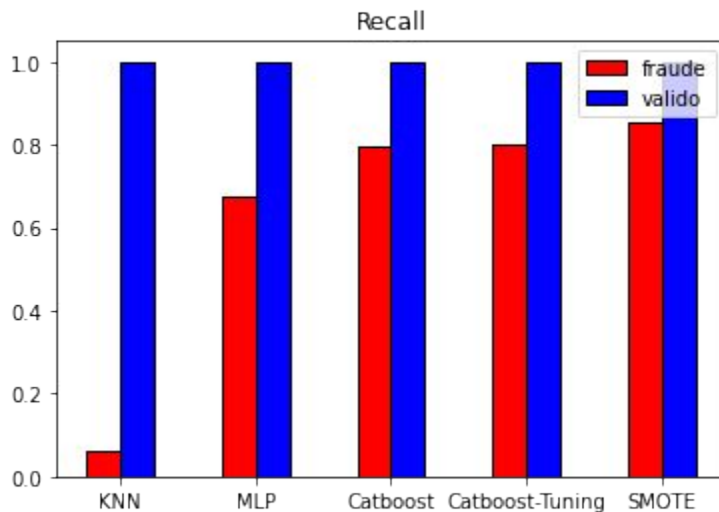


# Médias de Métricas por Modelo



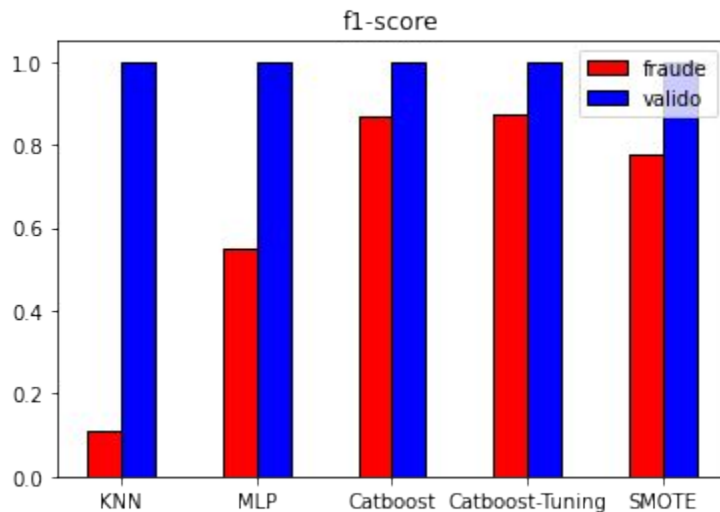


# Médias de Métricas por Modelo





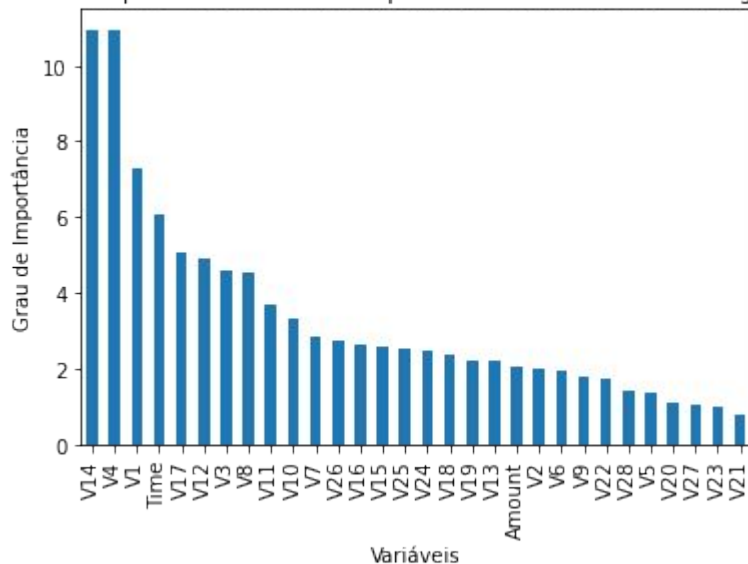
# Médias de Métricas por Modelo





# Importância de Variáveis (CatBoostClassifier-Tuning + SMOTE)

Análise de Importância de Variável para CatBoostClassifier-Tuning + SMOTE





# Conclusão

- Classificador CatBoostClassifier foi o que atingiu as melhores métricas gerais e para a Classe 1 (Fraudes).
- Apesar de pouco, foi possível obter desempenho a partir do tuning do classificador CatBoostClassifier.
- Utilização da técnica de oversampling SMOTE contribuiu para a melhor classificação da transacções fraudulentas.
- De forma geral, as técnicas empregadas cooperaram com o objetivo principal do trabalho, aumentar a taxa de Verdadeiros positivos para Classe 1, Fraudes.





# Obrigado

CatBoost. Catboostclassifier.

[https://catboost.ai/en/docs/concepts/python-reference\\_catboostclassifier.](https://catboost.ai/en/docs/concepts/python-reference_catboostclassifier)

imbalanced-learn developers. Smote.

[https://imbalanced-learn.org/stable/references/generated/imblearn.over\\_sampling.SMOTE.html.](https://imbalanced-learn.org/stable/references/generated/imblearn.over_sampling.SMOTE.html)

scikit-learn developers. sklearn.model selection.kfold.

[https://scikit-learn.org/stable/modules/generated/sklearn.model\\_selection.KFold.html#examples-using-sklearn-model-selection-kfold.](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html#examples-using-sklearn-model-selection-kfold)

scikit-learn developers. sklearn.neighbors.kneighborsclassifier.

[https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html.](https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html)

scikit-learn developers. sklearn.neighbors.mlpclassifier.

[https://scikit-learn.org/stable/modules/generated/sklearn.neural\\_network.MLPClassifier.html.](https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html)

ULB, M. L. G. Credit card fraud detection.

[https://www.kaggle.com/mlg-ulb/creditcardfraud/.](https://www.kaggle.com/mlg-ulb/creditcardfraud/)