

Fluxo de Trabalho

- 1° Notebook
- 2° Notebook
- 3° Notebook
- GIS

Limpeza dos dados

Análise exploratória

Descrição Univariada

Descrição Bivariada

Aquisição dos sensores

Bandas Landsat (5 features)

Grids Geofísicos (8 features)

Magnetométricos (1 feature)

14 features

Preparação dos sensores

Divisão treino/teste

Amostragem aleatória de 150 instâncias por unidade litoestratigráfica

Caso número de instâncias de uma unidade < 100, 70% dessa unidade será aleatoriamente amostrada

Conjunto de treino

Escalonamento das features

Estandardização

Redução da dimensionalidade

PCA

Superamostragem

SMOTE

Pré-processamento

Classificadores

Multinomial Logistic Regression

Decision Tree

Gaussian Naive Bayes

K-Nearest Neighbor

Support Vector Machines

Random Forest

XGBoost

Multilayer Perceptron

Otimização dos hiperparâmetros

Grid Search

Random Search

Estimação do erro de generalização

Validação cruzada K-fold

Avaliação dos modelos

Matriz de confusão

Acurácia

F1-Score

Recall

Precisão

Entropia Cruzada

Métricas

Desempenho do estimador do erro

Explicação do modelo

Pós-processamento