

- GIS
- notebook1
- notebook2

Fluxo de trabalho

1ª etapa

Aquisição dos sensores

- SRTM (1 feature)
- Bandas Landsat 8 (9 features)
- Bandas Landsat 7 (3 features)
- Radiométricos (7 features)
- Magnetométricos (1 feature)

21 features

Preparação dos sensores

- +espg = 31983
- +proj = UTM
- +zona = 23S
- +datum = SIRGAS2000
- +elp = GRS 1980
- +resol = 62.5 x 62.5 m

Sistema de referência

2ª etapa

Limpeza dos dados

Truncamento dos canais radiométricos

Análise exploratória

- Univariada
- Bivariada

3ª etapa

Divisão treino/teste

- Amostragem aleatória de 100 instâncias por unidade litoestratigráfica
- Caso número de instâncias de uma unidade < 100, 70% dessa unidade será aleatoriamente amostrada

Treino

Escalonamento das features

StandardScaler

Redução da dimensionalidade

PCA

Apenas bandas Landsat 8

Seleção das features

- Teste f ANOVA
- Informação mútua
- Gini / Entropia
- Regressão Logística

Features numéricas x target discreto

A partir de modelos

Superamostragem

SMOTE

Modelagem dos sensores

- Multinomial Logistic Regression
- Ridge Classifier
- Gaussian Naive Bayes
- K-Nearest Neighbor
- Support Vector Machines
- Random Forest
- Xgboost
- Artificial Neural Networks

Validação dos modelos

Métricas

- Matriz de confusão
- F1-score
- Precisão
- Revocação
- Acurácia

weighted e macro

Validação cruzada

Validação cruzada estratificada (5 folds)

Otimização dos modelos

- RandomSearch
- GridSearch

Erro da validação x Erro do teste

4ª etapa

Pós-processamento

Redução de ruídos

Filtro convolucional

Confecção dos mapas preditivos finais

Mapa geológico preditivo