

- 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)

22 features

Preparação dos sensores

- Reprojeção
- Padronização da resolução
- Padronização da extensão
- Clip
- Junção espacial

2ª etapa

Limpeza dos dados

Truncamento dos canais radiométricos

Análise exploratória

- Visualização espacial
- Descrição univariada
- Descrição 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

Conjunto de treino

Escalonamento das features

Estandardização

Redução da dimensionalidade

PCA

Seleção das features

Teste F ANOVA

Superamostragem

SMOTE

Classificadores

- Multinomial Logistic Regression
- Decision Tree
- Gaussian Naive Bayes
- K-Nearest Neighbor
- Support Vector Machines
- Random Forest
- XGBoost
- Multilayer Perceptron

Pipeline

Tuning dos hiperparâmetros

- Grid Search
- Random Search

Validação cruzada

Validação cruzada K-fold

Avaliação dos modelos

- Matriz de confusão
- Acurácia
- F1-Score
- Recall
- Precisão

Métricas

Avaliação do estimador do erro

Mapas de probabilidade

Mapas de entropia cruzada

Random Forest e XGBoost

4ª etapa

Pós-processamento

Redução de ruídos

Confecção dos mapas preditivos finais

Mapa geológico preditivo