

CURSO SISTEMAS DE INFORMAÇÃO

MINERAÇÃO DE DADOS

REGRESSÃO UTILIZANDO O PYCARET

Essa atividade tem por objetivo aplicar os conhecimentos de machine learning supervisionado a biblioteca PyCaret para resolver uma tarefa de **regressão** com um conjunto **de dados reais**. Dessa forma, é preciso que você explique todos os atributos utilizados do dataset.

Siga os seguintes passos:

- Importe e explore os dados escolhidos.
- Configure o ambiente do PyCaret para a tarefa de classificação, definindo a variável alvo.
- Compare os modelos disponíveis e selecione o melhor desempenho.
- Realize o ajuste de hiperparâmetros com `tune_model`.
- Avalie o modelo final, avalie as métricas mais importantes utilizando SHAP.
- Interprete os resultados: comente a performance do modelo, principais variáveis e possíveis limitações dos dados utilizados.
- Justifique a possibilidade desse modelo treinado ser utilizado ou não em um ambiente de produção.

Entrega:

Código Python comentado (.ipynb). Especifique dentro do código, a origem dos dados e a informação contida em cada atributo.

Na próxima aula, cada aluno deverá explicar o seu trabalho, inclusive todos os processos realizados pelo PyCaret.

Cada aluno apresentará o seu modelo treinado conforme a relação a seguir:

ALUNO	DATASET	PROBLEMA
ALFREDO DE SOUZA AGUIAR NETO	https://www.kaggle.com/datasets/aayushchou/hospital-length-of-stay-dataset-microsoft	Predição de tempo de internação hospitalar
ÁLVARO CALEBE SOUZA FERREIRA	https://www.kaggle.com/datasets/alamshihab075/mental-health-dataset	Predizer produtividade de autoavaliada
ANA BEATRIZ MARQUES MOREIRA	Ana Beatriz Marques Moreira - https://www.openml.org/search?type=data&status=active&qualities.NumberOfClasses=lte_1&id=46283	Prever o consumo de energia de eletrodomésticos numa residência
CAMILLE GOMES DE MENEZES (VICTOR HUGO)	https://www.kaggle.com/datasets/mohamedasak/imdb-top-250-movies	IMBD Top 250 Movies
CHARLES ABRAÃO PANTOJA	https://www.kaggle.com/datasets/devansodariya/student-performance-data	prever a performance dos estudantes

SANTOS		
FELIPE FERREIR A SOUSA	https://www.kaggle.com/code/leomauro/regress-o-pre-o-de-venda-da-casa	Preço de venda da Casa
FILIPPE DIAS BARRETO	https://www.kaggle.com/datasets/sudalairajkumar/daily-temperature-of-major-cities	Prever a temperatur a média diária de uma cidade
GABRIEL ALVES DA SILVA	https://www.kaggle.com/code/gustavolenin/preco-de-combustiveis-2004-a-2020	preço de combustive is - 2004 a 2020
GUILHER ME SILVEIRA FIGUEIRA	https://www.kaggle.com/datasets/toramky/automobile-dataset	Prever o valor de mercado de automóveis
HEMILLY CHRISTI NNE SILVA PEREIRA	https://www.kaggle.com/datasets/rudraprasadbhuyan/spotify-dataset-for-ml-practice	Spotify Dataset for ML Practice
HENRIQ UE PONTES BARROS AMORIM	Henrique Pontes - https://www.kaggle.com/datasets/spittman1248/cdc-data-nutrition-physical-activity-obesity	Compreend er como hábitos alimentares
ISAC PEREIRA DA SILVA	https://www.kaggle.com/datasets/uciml/red-wine-quality-cortez-et-al-2009	Prever o nível de qualidade de um vinho
JOANA		



UNITINS
Universidade Estadual do Tocantins

TOCANTINS
GOVERNO DO ESTADO



LETÍCIA AZEVEDO DO NASCIME NTO SOARES	https://www.kaggle.com/datasets/lunthu/steam-monthly-average-players?resource=download	
JOÃO VICTOR DE OLIVEIRA CARNEIR O	https://www.kaggle.com/datasets/safaruzzamanshovo/graduate-admission-dataset	Graduate Admission Dataset
JOÃO VICTOR FERNAN DES CORDEIR O MARTINS		
JOÃO VICTOR PÓVOA FRANÇA	https://www.kaggle.com/datasets/martj42/international-football-results-from-1872-to-2017	Internation al football results from 1872 to 2025
JÚLIA PEREIRA GONÇAL VES	https://www.kaggle.com/datasets/wickedbotz/teste-desmatamento-amazonia	Desmatam ento da Amazônia
KAIO MACEDO MACHAD O		
KATRINA	https://www.kaggle.com/datasets/nikhil7280/student-performance-	prever o



UNITINS
Universidade Estadual do Tocantins

TOCANTINS
GOVERNO DO ESTADO



DE MORAIS SANTOS	multiple-linear-regression?utm_source=chatgpt.com	índice de desempenh o de estudantes
LEANDR O GUIMARÃ ES GARCIA	https://www.kaggle.com/datasets/dongedong/beer-consumption-sao-paulo	Beer Consumpti on
LUCAS ARAUJO DE CASTRO	https://www.kaggle.com/datasets/dhruv1999/student-performance-dataset	Prever a nota final do estudante
LUCAS DANIEL RODRIG UES DOS SANTOS	LUCAS DANIEL RODRIGUES DOS SANTOS Dataset: NFL Stats 2012-2024 https://www.kaggle.com/datasets/philiphyde1/nfl-stats-1999-2022	Prever quais times jogariam a pós temporada da NFL.
LUCAS HENDER SON VIEIRA CUNHA	https://www.kaggle.com/datasets/denkuznetz/taxi-price-prediction	Problema: Prever o preço de corridas de táxi com base em diversos fatores, como distância, horário, condições de tráfego e outros.

LUÍS EDUARD O NOLETO DE MELO		
MARCEL O ALVES LEITE	https://www.kaggle.com/datasets/rhythmcam/pycaret-regression-auto-model/data	Prever a quantidade de itens vendidos
MATEUS TOSHIO COLI SHIBATA		
OTAVIO NARDINI DE FARIA	https://www.kaggle.com/datasets/rehan497/worldwide-temperature-trends-19612024	
VICTOR ALVES DOS REIS DIAS	Victor Alves Conjunto de dados Auto MPG https://www.openml.org/search?type=data&status=active&id=42372	