

	Instituto Federal de Educação Ciência e Tecnologia – Sudeste de Minas Gerais		
	Disciplina: Machine Learning		Data: 17/05/2022
	Professor: Gustavo Willam Pereira		
	Trabalho Prático I	Grupo: 02 Alunos	Valor: 30,0 pontos
	Forma de Entrega: SIGAA		Data de Entrega: 07/06/2022

O trabalho consiste na elaboração de um algoritmo de Machine Learning para problema de Regressão (Regressão Linear Simples, Regressão Linear Múltipla, Support Vector Machine, Decision Tree, Random Forest).

Realize Análise Exploratório dos Dados (EDA)  
 Implemente técnicas de Eliminação de Outliers  
 Seleção de Variáveis (*features*)  
 Normalize/Padronize os dados de for necessário.  
 Realize o tratamento de dados categóricos.  
 Implemente otimização de Parâmetros

A proposta do algoritmo de Machine Learning é participar da seguinte competição do Kaggle: <https://www.kaggle.com/competitions/house-prices-advanced-regression-techniques>

Os arquivos contendo os conjuntos de dados de treino e teste, explicação do dataset, e modelo de submissão estão na plataforma Kaggle ou podem ser acessados através do GitHub: <https://github.com/gustavowillam/ML/tree/main/trabalhos/2022/Trab1>

Critérios para Avaliação:

- 1) (15,0 pontos) Submeter o trabalho no Kaggle e verificar sua pontuação.
  - a. São aceitos até 10 submissões diárias. Realize simulações, implemente soluções distintas com objetivo de melhorar seu Ranking de pontuação.
- 2) (15,0 pontos) Apresentação do trabalho: O trabalho deverá ser apresentado em sala de aula. Deverá ser explicado como foi implementado cada técnica: EDA, Eliminação de Outliers, Escalonamento dos dados, técnicas utilizadas para tratar dados categóricos, otimização de parâmetros, seleção de *features*, etc... Deverá ser executado o código, preferencialmente no Google Colaboratory.
- 3) Ponto Extra:
  - a. 15 pontos (1º colocado do Ranking interno da sala de aula)
  - b. 10 pontos (2º colocado do Ranking interno da sala de aula)
  - c. 05 pontos (3º colocado do Ranking interno da sala de aula)