



Instituto Federal de Educação Ciência e Tecnologia – Sudeste de Minas Gerais

Disciplina: Visualização de Dados

Data: 10/05/2023

Aluno: _____

Valor: 3,0 pontos

Data da Entrega: 31/05/2023

Grupo: 02 alunos

Entrega: SIGAA

Considere o banco de dados SQLite3: StudentsPerformance.db

O objetivo deste trabalho é implementar um algoritmo de Machine Learning (ML) ou de Rede Neural (RNA) para realizar a predição do valor da Nota de Matemática (math score).

Considere como variáveis independentes (X):

- a) gender (sexo)
- b) race_ethnicity (raça/etnia)
- c) parental level of education (nível de escolaridade dos pais)
- d) lunch (almoço)
- e) test preparation course (curso preparatório para testes)
- f) reading score (pontuação de leitura)
- g) writing score (pontuação de escrita)

Considere como variável dependente (y):

math score (pontuação de matemática)

1) Implemente uma solução de ML “do zero” . Utilize um modelo de ML de sua preferência, ou Pycaret, ou Redes Neurais.

2) Seu modelo deverá consumir uma base de dados em formato .csv, que será gerada pelo Apache Airflow que ficará responsável por “orquestrar” os dados em produção com os dados a serem consumidos pelo seu modelo de ML.

3) Implemente uma interface gráfica (data app) utilizando o framework Streamlit para testar sua aplicação localmente.

4) Hospede sua aplicação no GitHub e através do Streamlit share, realize o deploy da sua aplicação.