

# Aprendizado de máquina- atividade discente supervisionada 2

Prof. Mozart Hasse

## **LEIA ATENTAMENTE TODAS AS INSTRUÇÕES ATÉ O FINAL DA ÚLTIMA PÁGINA. CADA PALAVRA CONTA!**

Use os microdados do ENEM de 2022 do estado do Paraná. Os nomes dos campos e seus significados podem ser baixados do site do INEP:

<https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem>

A base já foi filtrada e pré-tratada pelo professor. Ajustes adicionais poderão ser necessários dependendo da arquitetura escolhida.

O objetivo da sua equipe é PREVER A NOTA DE REDAÇÃO (NU\_NOTA\_REDACAO) a partir dos demais campos fornecidos pelo professor. Fica a seu critério ajustar formatos ou remover campos, **desde que o arquivo final gerado tenha as mesmas colunas dos dados de treinamento. Não é permitido enriquecer os dados com outras fontes ou usar campos excluídos pelo professor.**

**A solução deste problema DEVE OBRIGATORIAMENTE ser buscada usando redes neurais com execução 100% na máquina local.**

### Critérios de avaliação

**Organização e clareza do código: (30% da nota)**, incluindo comentários com a justificativa para as escolhas feitas na otimização.

**Precisão do resultado final: (70% da nota)** sobre a BASE DE TESTES FORNECIDA PELO PROFESSOR, os valores previstos serão comparados com os valores reais e a nota será dada pelo total dos quadrados das diferenças entre os valores previstos e os valores reais (conhecidos apenas pelo professor):

Raiz do erro quadrático médio	Nota aproximada
70000 ou acima	Zero (grosseiramente insuficiente)
15000 e abaixo	7,0
8500 ou inferior	10,0 (empate técnico com o professor :-D)

Cabe lembrar que compartilhar caminhos bem e mal sucedidos é permitido. O único cuidado é compartilhar também os **eventuais valores de parâmetros de configuração, que devem ter uma diferença de no mínimo 10% para mais ou para menos em pelo menos um parâmetro da implementação para cada equipe.**