**PROVA PRÁTICA PARA PESQUISADOR 2 - APRENDIZADO DE MÁQUINA**

Um dos métodos de *ensemble* (comitê) mais utilizados em competições é o método de *stacking*, que consiste no empilhamento de modelos. Nele, as entradas a partir da segunda camada de modelos passam a ser as inferências dos modelos anteriores.

**Pede-se nesta prova prática que o candidato treine um classificador binário incorporando o conceito de *stacking* em sua solução, e que esteja apto a justificar as escolhas que foram feitas ao longo do desenvolvimento.**

Para isso é fornecido um dataset para classificação binária com 800 observações para treinamento e 200 observações para teste, contando com 20 features. No entanto, apenas as observações de teste estão rotuladas. Pede-se que o candidato busque estratégias para rotular o conjunto de treinamento.

O conjunto de dados se encontra neste link: INSERIR LINK. Há três arquivos: *X\_train.csv*, *X\_test.csv*, e *y\_test.csv*.

O candidato é livre para escolher os métodos de solução. É desejável que a solução seja replicável e disponibilizada no github, com a instalação de um ambiente virtual ou uso de Docker, por exemplo.