

## **Utilização de grandes modelos de linguagem para geração de dados tabulares**

**Código do projeto:** PO01

**Responsável:** Profa. Patrícia Rufino Oliveira

**Linha de pesquisa:** Inteligência de Sistemas

**Número de vagas:** 2

### **Descrição geral**

A geração de amostras sintéticas que preservem as características originais de um determinado conjunto de dados é considerada um desafio. O objetivo deste projeto é explorar e avaliar grandes modelos de linguagem (LLM's) para a geração de dados tabulares. A hipótese é que tal abordagem pode ser aplicada para amenizar problemas como o desbalanceamento de classes e a existência de impurezas nos dados disponíveis para experimentos de aprendizado de máquina.