

Nomes: Bruno Pasquetti, Gabriel Brocco, Pedro H. De Bortoli e Rafael Klein

Concepção e Planejamento

Nesta fase, define-se o problema a ser resolvido: prever a evasão de alunos em um curso online. Primeiro é analisado se o problema pode ou não ser resolvido com Machine Learning. Se puder ser resolvido, são estabelecidos os objetivos do projeto, como reduzir a evasão por meio de intervenções baseadas em previsões, além da compreensão do contexto educacional e do impacto social da evasão. Também são escolhidas as métricas de avaliação, as ferramentas a serem utilizadas e a divisão das responsabilidades da equipe.

Experimentação

Consiste na análise e exploração dos dados, incluindo estatísticas descritivas, visualização e identificação de padrões relevantes. Nessa etapa, são testados diferentes modelos de Machine Learning, definidos para o objetivo que espera-se ser alcançado, ajustados seus hiperparâmetros e avaliados com base nas métricas definidas. O foco é descobrir quais técnicas têm maior poder preditivo para o problema de evasão.

Criação de Pipelines

Após definir os melhores métodos, esta fase automatiza o fluxo de trabalho: tratamento de dados, transformação de variáveis, treinamento e validação dos modelos. O objetivo é garantir organização e eficiência no processo de modelagem, facilitando ajustes futuros e integrando todo o processo de ponta a ponta de forma modular e escalável.

Produção

Nesta etapa, o modelo final é treinado com todos os dados disponíveis e exportado para ser utilizado em ambiente real. Documentar a pipeline e os requisitos para uso futuro. Técnicas de interpretabilidade (XAI) são aplicadas para tornar o modelo mais transparente, auxiliando na tomada de decisão pela equipe do curso online de onde os dados provêm.