PUC MG - TCC CIÊNCIA DE DADOS E BIG DATA

ETAPAS E RESULTADOS



DEFINIÇÃO DE PROBLEMA

Escolher um tema presente em meu dia a dia fez toda a diferença e agregou valor especial ao desenvolvimento do trabalho. Identificar clientes propensos a reclamar no Bacen e trabalhar de forma proativa, para que isso não aconteça!

COLETA DE DADOS

Esta etapa foi um desafío por si só! O ambiente onde os dados são mantidos é extremamente controlado (como não podía deixar de ser), mas consegui ter acesso a informações muito relevantes e que se mostraram assertivas (sempre adotando todas as precauções para que a segurança dos dados fosse preservada).





PRÉ PROCESSAMENTO

Já utilizei a palavra "desafio"?

Aqui, sem dúvida, foi onde grande parte do esforço do trabalho foi concentrado. Ajustar tudo, até que os dados pudessem, efetivamente, ser utilizados, foi um processo enriquecedor. Muito aprendizado foi produzido aqui.

ANÁLISE E EXPLORAÇÃO

Muitos "insights" interessantes puderam ser extraídos durante esta fase. A qualidade dos dados, como eles se apresentaram após todo o pré-processamento, disparidades, características importantes (como o desbalanceamento), etc. Não adiantava ter pressa....





CRIAÇÃO DO MODELO

Confesso que este foi o momento que mais gerou ansiedade: "o que o modelo vai me dizer?!", "qual algoritmo melhor se encaixa no problema?". Enfim, embora nem de longe a etapa mais trabalhosa, acaba sendo a que "materializa", de certa forma, todas as anteriores.

RESULTADOS

Muitas informações importantes para que os gestores responsáveis possam se debruçar sobre processos e produtos que acabaram se mostrando como mais recorrentes nas reclamações.

E, agora, um modelo de "Machine Learning" pode apoiar na tomada de decisão, identificando clientes que demandem maior nível de atenção.



