Diário de Desenvolvimento

**19/10/2024**

Eu tive problemas com o dataset que estou utilizando para o desenvolvimento do problema, no caso existem classes no dataset de algumas variáveis categóricas que não existem no dicionário de dados, portanto, elas não têm identificação. Eu tentei procurar resposta no site onde peguei o dataset e no Stack overflow, mas não encontrei nada. Eu extraí manualmente essas classes para trata-las, esse é o 1º obstáculo do projeto

Decidi ignorar classes que tem no dicionário, mas não existem no dataset

**05/11/2024**

Para melhor decidir como agrupar as features do dataset com o intuito de avaliar a correlação delas entre si em grupos e contra a variável target, eu precisava compreender melhor o dataset em si, por isso fui atrás do artigo que criou o dataset nomeado **Early prediction of student's performance in higher education: a case study.** Descobri se tratar de um capítulo de um livro chamado **Trends and Applications in Information Systems and Technologies,** custei para encontrar um PDF gratuito na web, chegando até cogitar enviar mensagens aos autores do artigo para solicitar uma cópia, mas encontrei e continuei meu trabalho a partir desse ponto.