



COMUNIDADE DS



Google Cloud Platform

Data Science no GCP

Projeto do Aluno

Agenda

1. Problema de negócio
2. Produto de dados esperado
3. Dados
4. Dicas
5. Finalizar



— Problema de negócio

Você recebeu a tarefa de criar um modelo de machine learning para auxiliar uma construtora a medir a resistência do concreto, pois eles dependem dessa questão para construir os seus projetos, e a equipe de engenheiros está interessada em prever a resistência do concreto, que é a parte mais importante da construção civil, dado algumas características que eles têm disponíveis.



— Produto de dados esperado

Como ainda existe uma certa dificuldade da equipe em confiar em um algoritmo de machine learning, foi sugerido para esse momento ter uma POC, que será consumido e validado pela equipe de engenheiros, essa POC vai ser um dashboard que vai consumir as previsões feitas pelo modelo e vai ter gráficos onde a equipe poderá validar se a solução realmente está de acordo com o que o negócio espera.



Dados

Os dados estão armazenados no:

- https://drive.google.com/file/d/1kzsPaYtSgOi3RIg6ikvrBIKy_aH0ONDIG/view?usp=sharing

Sobre os dados:

- 9 colunas, sendo 8 colunas variáveis independentes e a outra é a variável dependente
- Variável que você deve prever: csMPa



COMUNIDADE DS

*Todos esses valores estão presentes por cada kilo em uma mistura metros cúbicos

Dados

Nome da coluna	Descrição
cement*	Cimento
slag*	Escória granulada de alto forno
flyash*	A cinza volante é um resíduo finamente dividido da combustão do carvão pulverizado
water*	Água
superplasticizer*	Superplastificantes
coarseaggregate*	Agregados grosseiros são quaisquer partículas maiores que 0,19 polegadas
fineaggregate*	Agregados finos geralmente consistem em areia natural ou brita com a maioria das partículas, passando por uma peneira de 3/8 de polegada.
age	Idade em dias (1 até 365)



COMUNIDADE DS

Dicas

Algumas dicas

- Entenda o que problema ser resolvida
 - Que tipo de problema estamos falando?
- Planeje a solução
 - Crie um esqueleto do projeto
 - Desenhe o dashboard antes de criar
 - Entenda os dados
- Desenvolva o projeto
 - Vertex Workbench
 - Local
- Organize bem a estrutura do seu projeto
- Como mostrar uma boa predição
- Como ganhar a confiança da equipe com o algoritmo?
- Vertex ou BigQuery?
- Faça EDA
- Crie um diagrama de fluxos



COMUNIDADE DS

Finalizando

- Precisa de minha ajuda? Manda a sua dúvida no canal *gcp-data-science* lá no Discord, bom que pode ajudar outras pessoas!

Vou avaliar alguns projetos, podemos até fazer em aula, mas preciso que envie o projeto para:

- matheus.oliveira@comunidadeds.com
 - Conteúdo: github com a sua solução, dashboard em qualquer ferramenta e tudo que fizer além
 - Não esqueça de me avisar no discord sobre o envio!
 - Podemos programar para uma apresentação

OBRIGADO!



COMUNIDADE DS