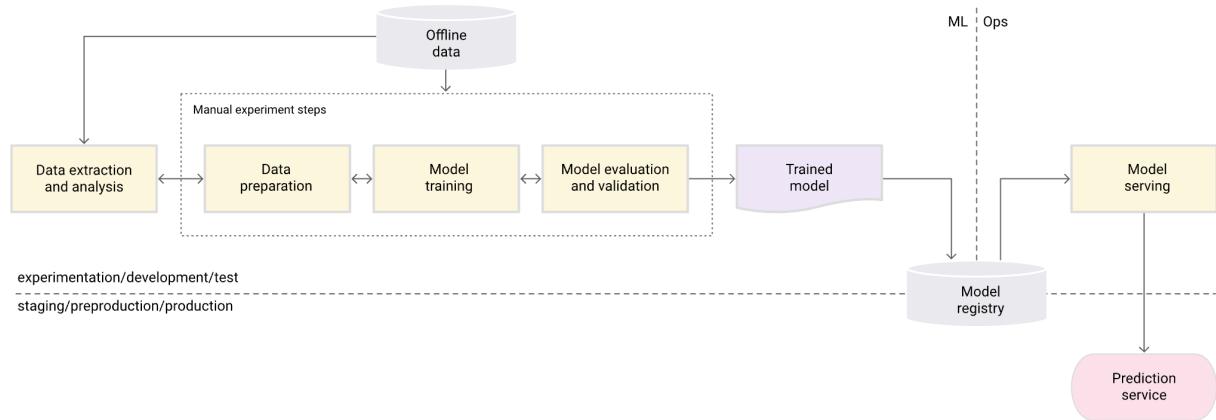


Projeto Final Tutorial MLE

Proposta de Projeto

Neste projeto final você deve construir um [pipeline de MLOps de nível 0](#) para resolver um case de negócios.



Fonte: https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning?hl=pt-br#mlops_level_0_manual_process

O que você vai entregar?

Ao final do seu projeto você deve entregar:

- Um repositório no Github/Gitlab com todas as implementações;
- Um vídeo de até 15 minutos apresentando seu projeto. Nele você deve responder às seguintes perguntas:
 1. Descreva como obteve os dados.
 2. Apresente quais foram as análises que você realizou sobre eles e o que concluiu com elas?
 3. Como os dados foram selecionados e preparados para o treinamento?
 4. Qual modelo você escolheu? Por que?
 5. Como você definiu que esse modelo estava adequado para servir?
 6. Como você construiu o serviço do modelo? Por que você escolheu dessa forma?
 7. Como você faria para monitorar este modelo em produção? O que você acompanharia?

Instruções gerais

- Você pode usar [Localstack](#) para construir a sua solução. Se preferir você pode usar algum provedor de cloud, você pode. No entanto, não reembolsaremos nenhum gasto durante o desenvolvimento do projeto final ou qualquer outra etapa do curso.

- Você pode usar ferramentas como o [Google Colab](#) para ter acesso à GPUs/TPUs durante o treinamento de forma gratuita;

Case sugerido

Em um mercado cada vez mais competitivo, a satisfação do cliente é um dos principais indicadores de sucesso para um e-commerce.

Entender como os clientes percebem os produtos e serviços oferecidos pode proporcionar insights valiosos para aprimorar a experiência de compra, otimizar o portfólio de produtos e orientar estratégias de marketing.

Uma abordagem eficiente para isso é a análise de sentimentos em avaliações de produtos, permitindo classificar opiniões como positivas ou negativas de forma automática e em escala.

Desafio

Você deve implementar um pipeline de MLOps de nível 0 para esse avaliador de sentimentos descrito no item anterior.

Dados

<https://www.kaggle.com/datasets/kritanjilijain/amazon-reviews/data>

Dica: este dataset é muito grande, talvez seja interessante usar apenas uma parte das amostras dele. Descreva como fez a seleção dos dados.

Cases alternativos

Caso queira atuar sobre algum outro problema, sinta-se livre para propor! Mas atente-se à algumas regras:

1. **Nunca** use dados do banco, use apenas dados para simular o seu problema. O Kaggle está cheio de bons datasets, procure um equivalente lá ou algum dataset aberto sobre o tema;
2. Você deve apresentar o seu case no começo do seu vídeo, nesse caso seu vídeo pode ter até **20 minutos** no total.