



Menu

Pré-processamento de dados

Traduzido automaticamente do Inglês

Este item inclui conteúdo que ainda não foi traduzido para o idioma de sua preferência.

O Apache Beam é um produto que oferece um modelo de programação unificado que permite implementar trabalhos de processamento de dados em lote e de fluxo contínuo em qualquer mecanismo de execução. Ele é ideal para o pré-processamento de dados!

Acesse <https://beam.apache.org/get-started/try-apache-beam/> para experimentar o Apache Beam em um Colab, para que o senhor possa ter uma ideia de como as APIs funcionam. Certifique-se de experimentá-lo em Python e Java usando as guias na parte superior.

Observação: O senhor pode clicar no botão **Run in Colab** abaixo do snippet de código para iniciar o Colab. Na barra de menu do Colab, clique em **Runtime > Change Runtime type** e selecione **Python 3** antes de executar as células de código. O senhor pode obter mais explicações sobre o exemplo do WordCount [aqui](#) e também pode usar o [Beam Programming Guide](#) para pesquisar qualquer um dos conceitos.

O senhor pode aprender sobre a transformação do TensorFlow aqui:

https://www.tensorflow.org/tfx/transform/get_started . Ele também usa pipelines no estilo do Beam, mas tem módulos otimizados para o pré-processamento de conjuntos de dados do Tensorflow.

Marcar como concluído

Curtir Não curtir Relatar problema

