

# 01\_pipelines\_intro\_kfp

- ▼ 개요 : Vertex AI Pipeline를 Kubeflow Pipelines SDK로 사용하는 법을 다룸
- ▼ 목적 : Vertex AI pipeline 정의 및 컴파일, pipeline run 서비스 계정 명시
- ▼ 행동
  1. 사용 컴포넌트 : Vertex AI, Cloud Storage, Vertex AI Workbench Notebook(or Local Specific Env)
  2. Installation : Workbench Notebook의 경우 실행해야하는 코드
  3. Before you begin : Google Cloud project 세팅
    - a. Region 세팅
    - b. Timestamp 세팅
    - c. Google Cloud 계정 인증 # Vertex AI Workbench에서는 필요없음
    - d. Cloud Storage 생성 및 세팅
    - e. Importing 및 환경변수(API\_ENDPOINT, PIPELINE\_ROOT) 세팅
  4. Initialize Vertex AI SDK for Python
    - a. Component Decorator로 간단한 python Docker Image를 사용하는 컴포넌트 선언
    - b. Pipeline 선언
  5. Compile the pipeline
    - a. 4번에서 수행한 코드들을 Compling
  6. Run the pipeline
    - a. PipelineJob으로 5번에서 Compile한 json파일을 run 수행
    - b. job.delete()로 삭제 가능
  7. Specifying a service account to use for a pipeline run : 서비스 계정 명시 가능
  8. Pipeline step caching : enable\_caching 명시 가능
  9. Using the Pipelines REST API : REST로 파이프라인 얻기?

## 10. Cleaning up : Google Cloud project 지우기

### ▼ 결과

#### 1. 해결한 점

- a. Vertex AI Workbench + Kubeflow pipeline 작성을 통해 기존 Kubeflow 작성법으로 수행할 수 있음
- b. 초기 Vertex AI + Cloud Storage Setting이 필요함

#### 2. 해결해야할 점

- a. 좀 더 MLOps Component 수준으로 깊은 Pipeline 요소들을 확인하고 싶음
- b. Docker Image를 활용하거나, packages\_to\_install 요소를 활용하거나 더 효율적인 방법을 찾고 싶음