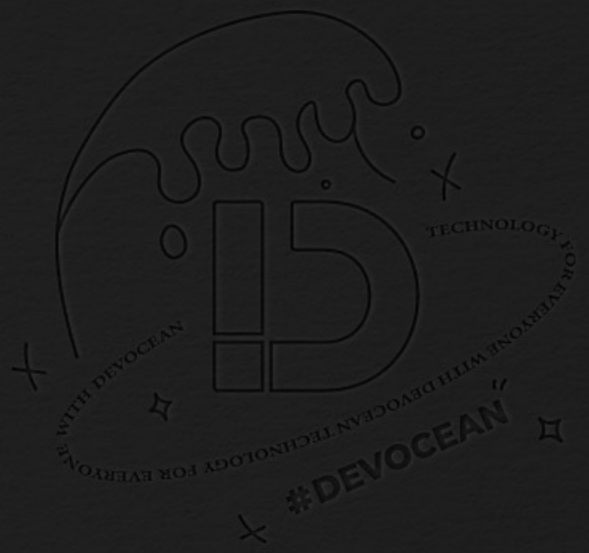
DEVOCEAN

Kubeflow 알아보기

김예은

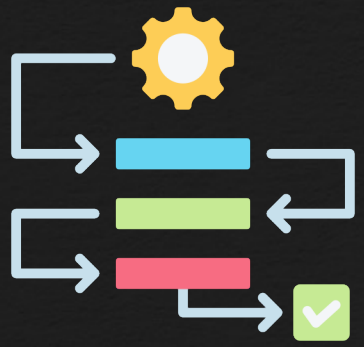
경험에 경험을 더하는 개발자, 김예은

목차



01 Kubeflow란?

Kubernetes



ML Workflow

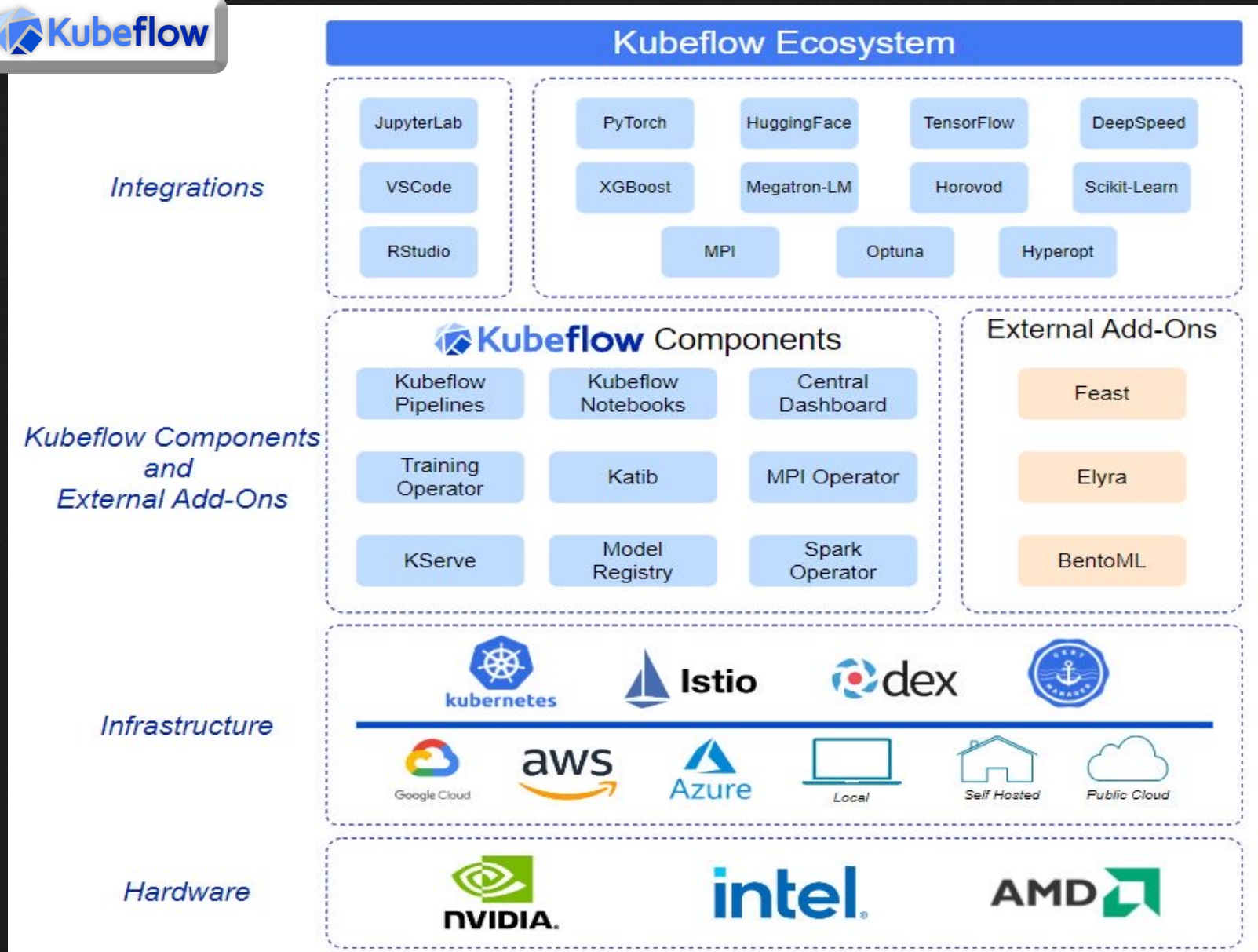
Kubeflow

- Kubernetes 환경에서 ML 라이프사이클의 각 단계를 다룰 수 있는 오픈소스 프로젝트
- 개발자, 데이터 사이언티스트, ML 엔지니어 등 모두가 모델 구축부터 테스트, 배포까지 ML 라이프사이클의 모든 부분을 처리할 수 있도록 서비스를 제공

01 Kubeflow란?

Why?

01 Kubeflow란?



02 Kubeflow Components

Pipeline
Kubeflow Pipeline

Notebooks
Kubeflow Notebooks

Dashboard
Kubeflow Central Dashboard

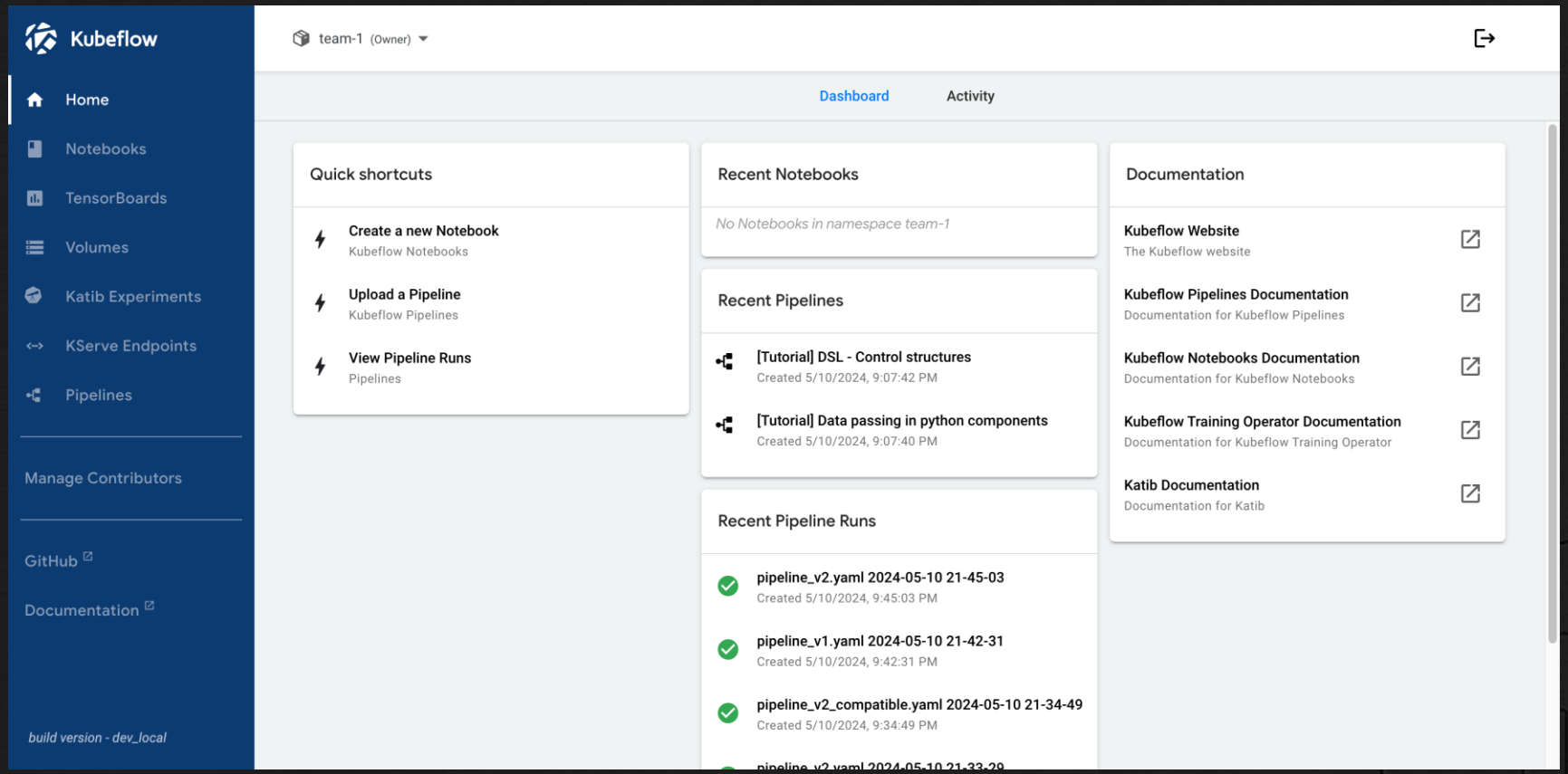
AutoML
Katib

Model Training
Kubeflow Training Operator

Model Serving
KServe

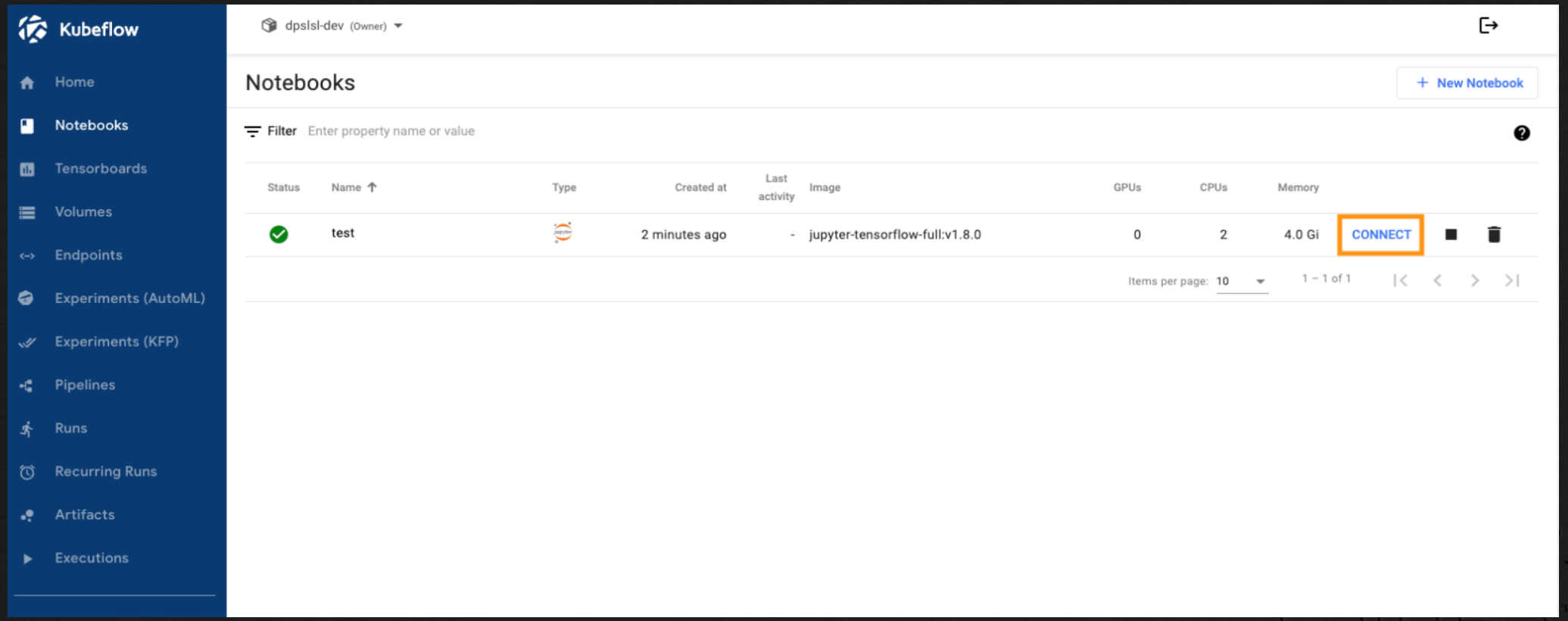
출처 : https://www.kubeflow.org/

02 Kubeflow Components – Central Dashboard



- 클러스터에서 실행되는 구성 요소의 실행 및 내용들을 확인할 수 있도록 웹 인터페이스를 제공
- 프로필을 이용하여, 사용자를 등록 및 삭제하여 관리하고 사용자에게 읽기, 쓰기 등의 권한을 부여

02 Kubeflow Components - Kubeflow Notebooks

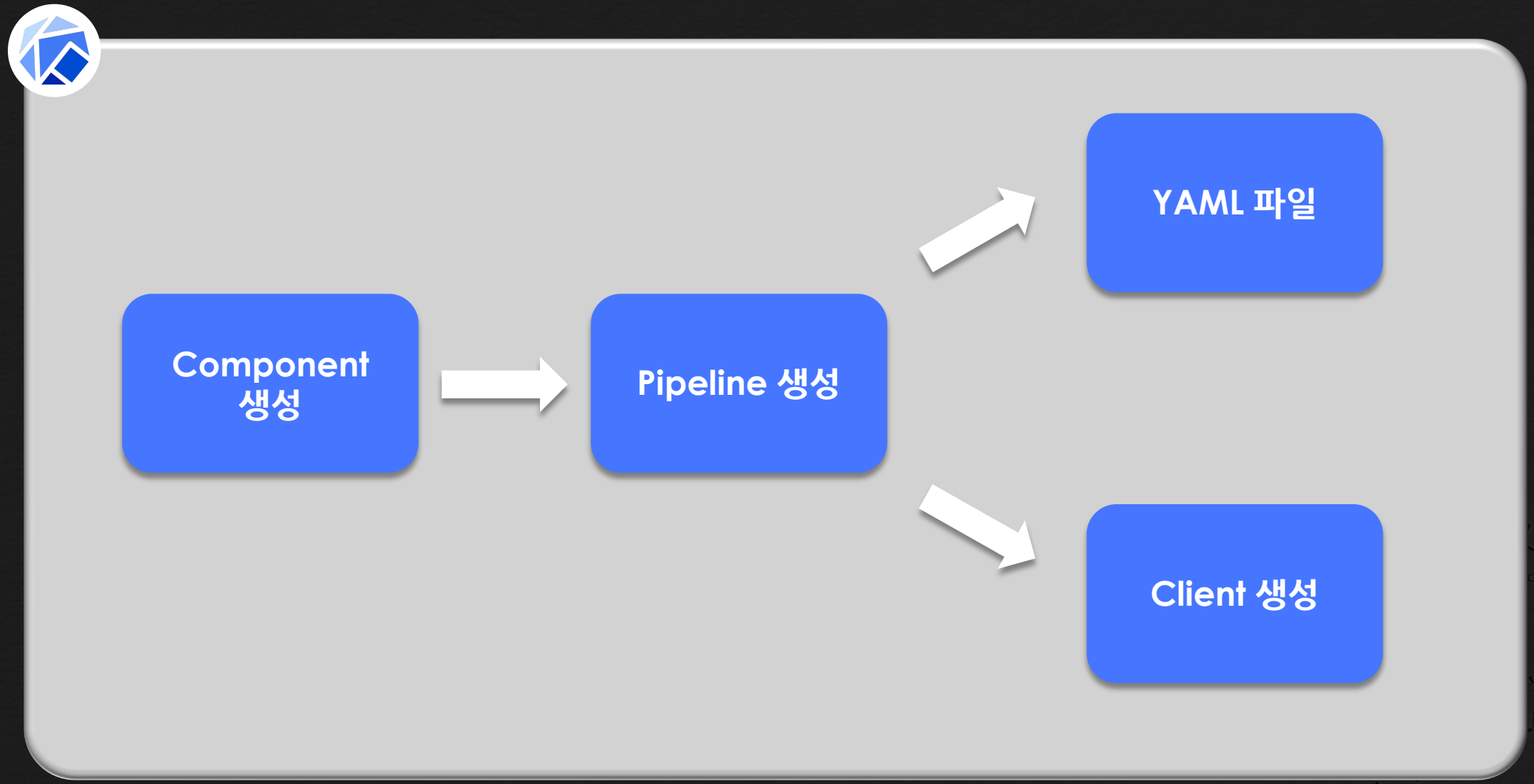


- Kubernetes 클러스터 내에서 웹 기반 개발 환경을 실행할 수 있는 방법을 제공
- JupyterLab , RStudio 및 Visual Studio Code를 기본적으로 지원(유형 및 패키지는 Docker 이미지로 선택 가능)

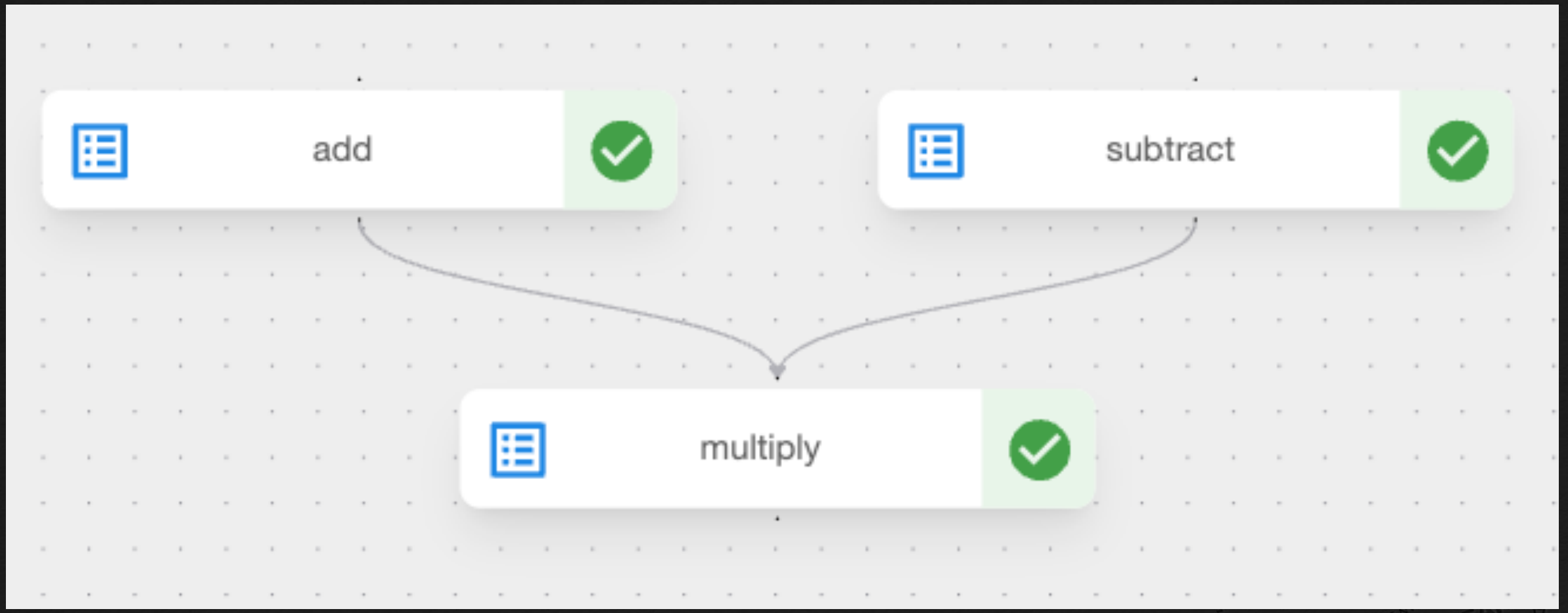
02 Kubeflow Components – Kubeflow Pipeline

- Kubernetes를 사용하여 이식과 확장이 가능한 ML 워크플로우를 구축, 배포하기 위한 플랫폼
- KFP Python SDK를 사용하거나, YAML 파일을 이용하여 누구나 쉽게 Pipeline 구축 가능
- 사용자가 커스텀하거나, 기존의 ML 구성 요소들을 활용해서 다양하게 사용 가능
- 파이프라인 정의, 실행, 실험 등을 쉽게 관리하고 시각화로 확인 가능

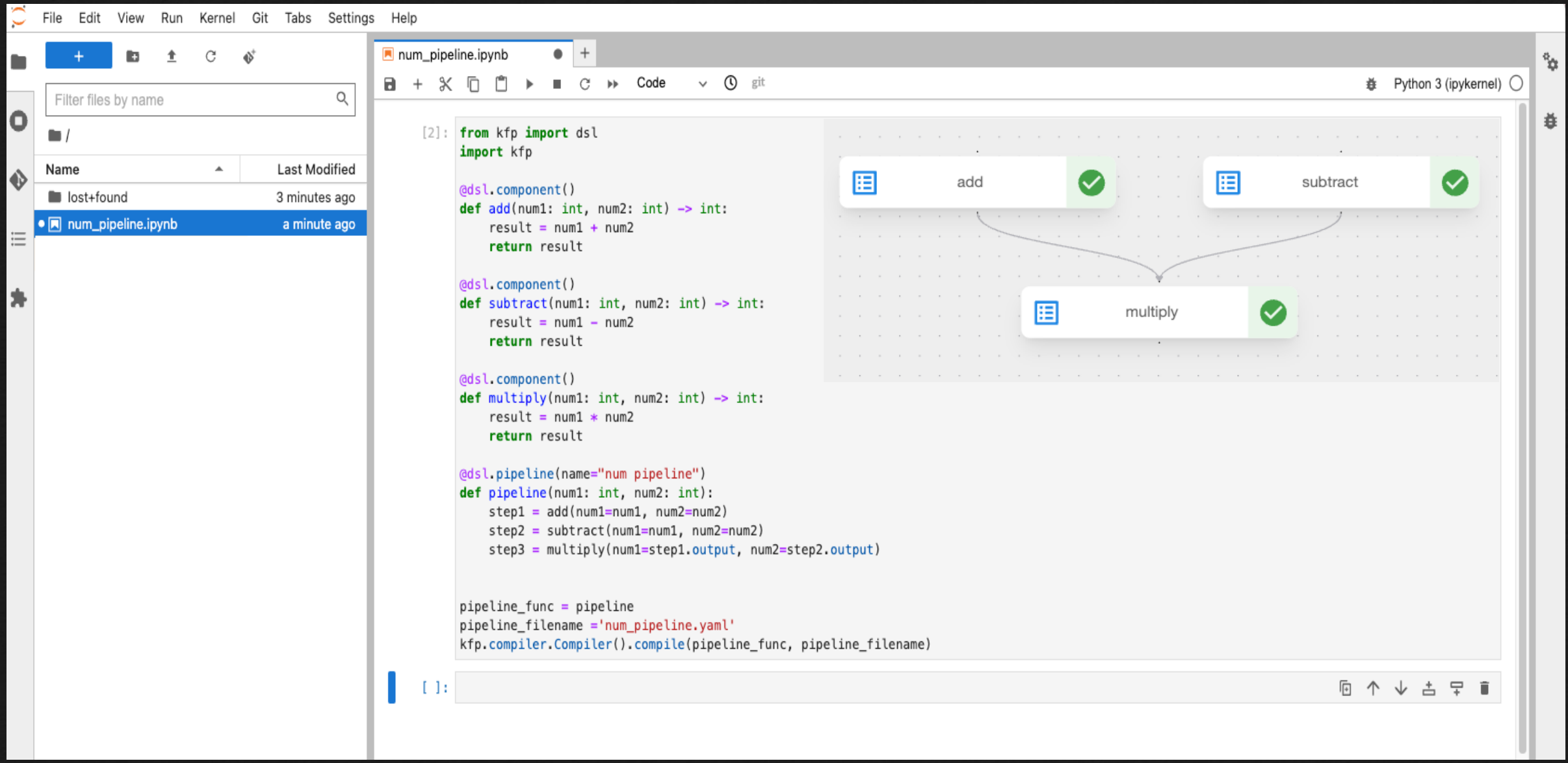
02 Kubeflow Components – Kubeflow Pipeline



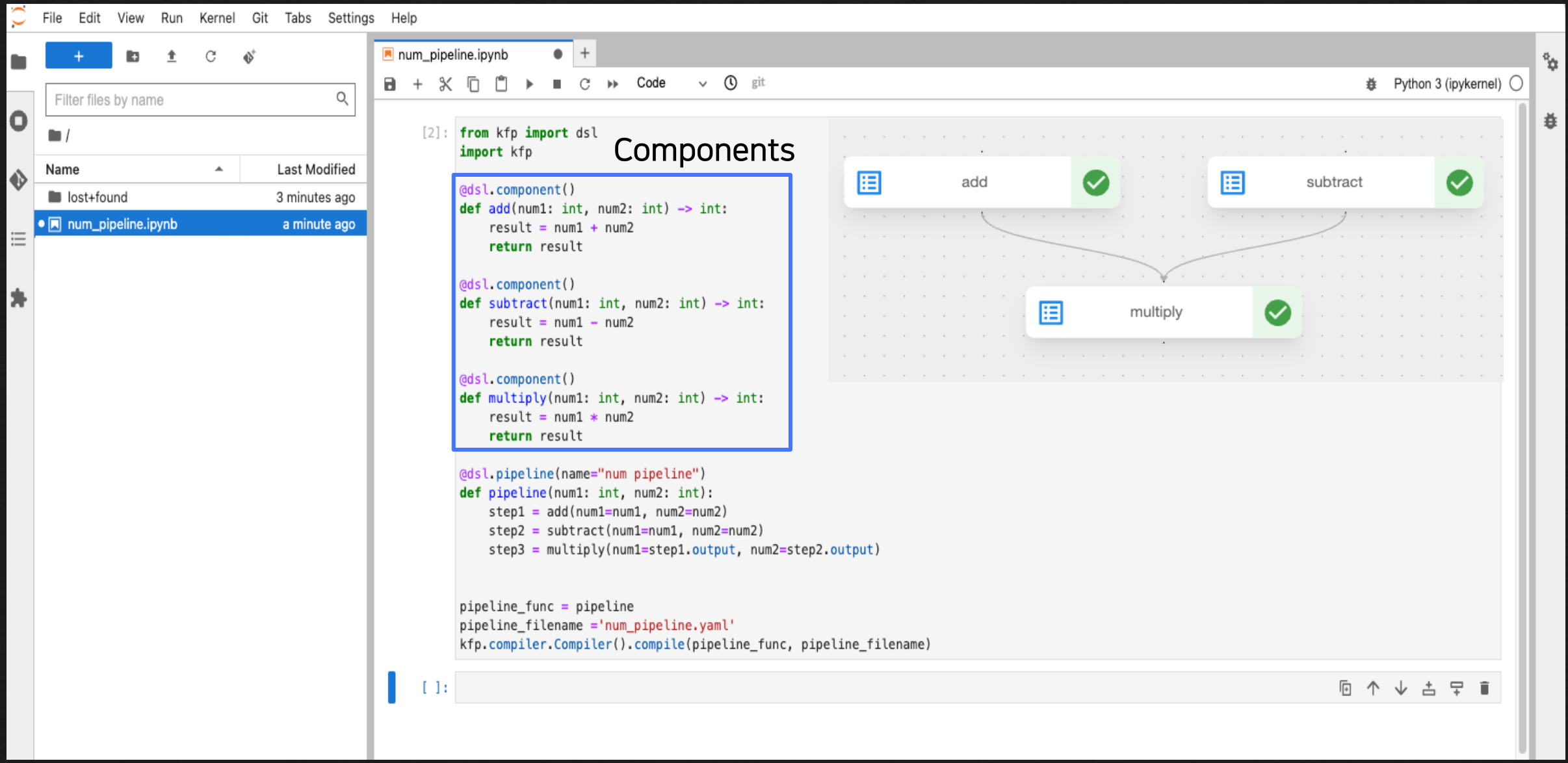
02 Kubeflow Components – Kubeflow Pipeline



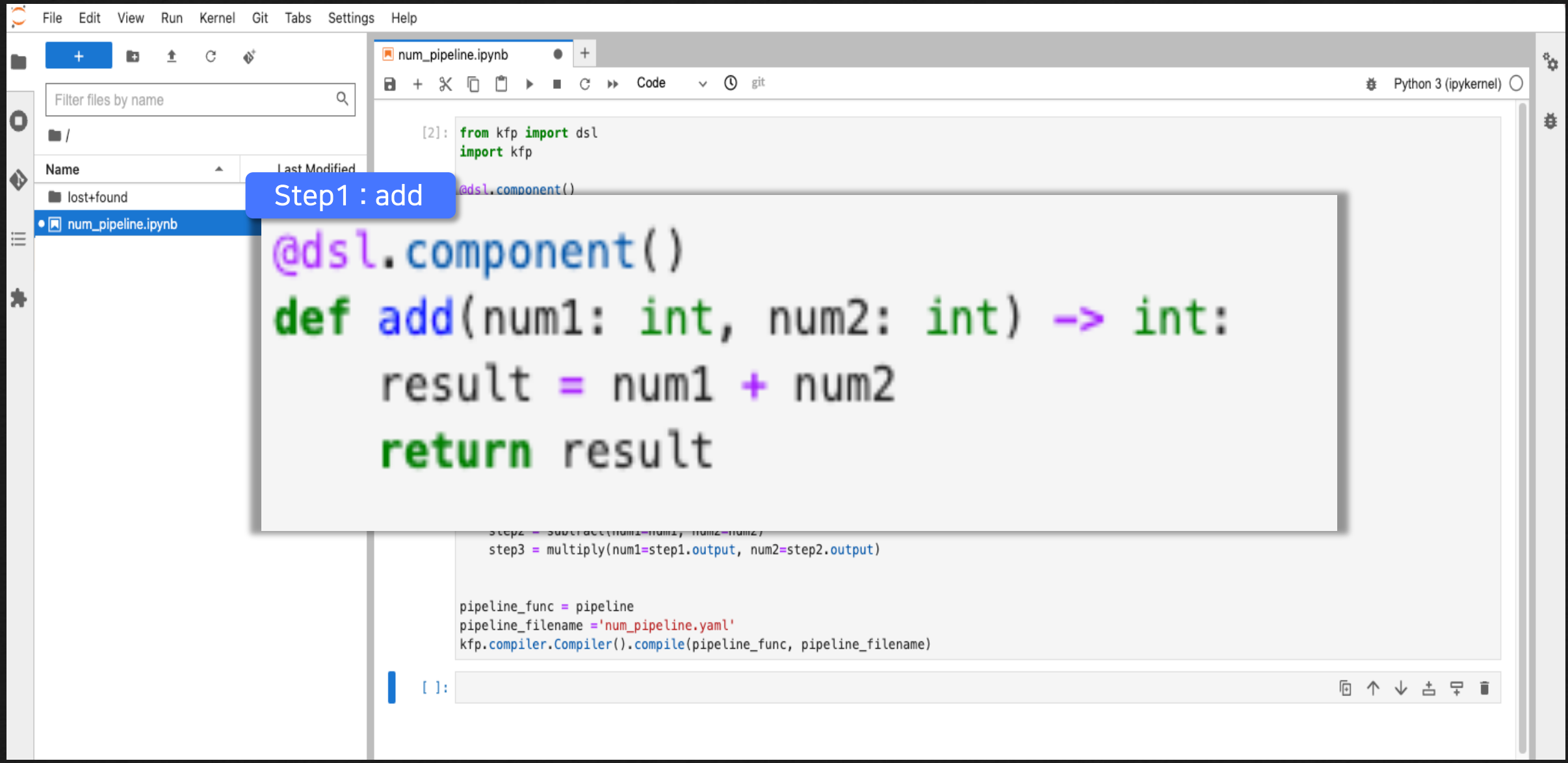
02 Kubeflow Components – Kubeflow Pipeline



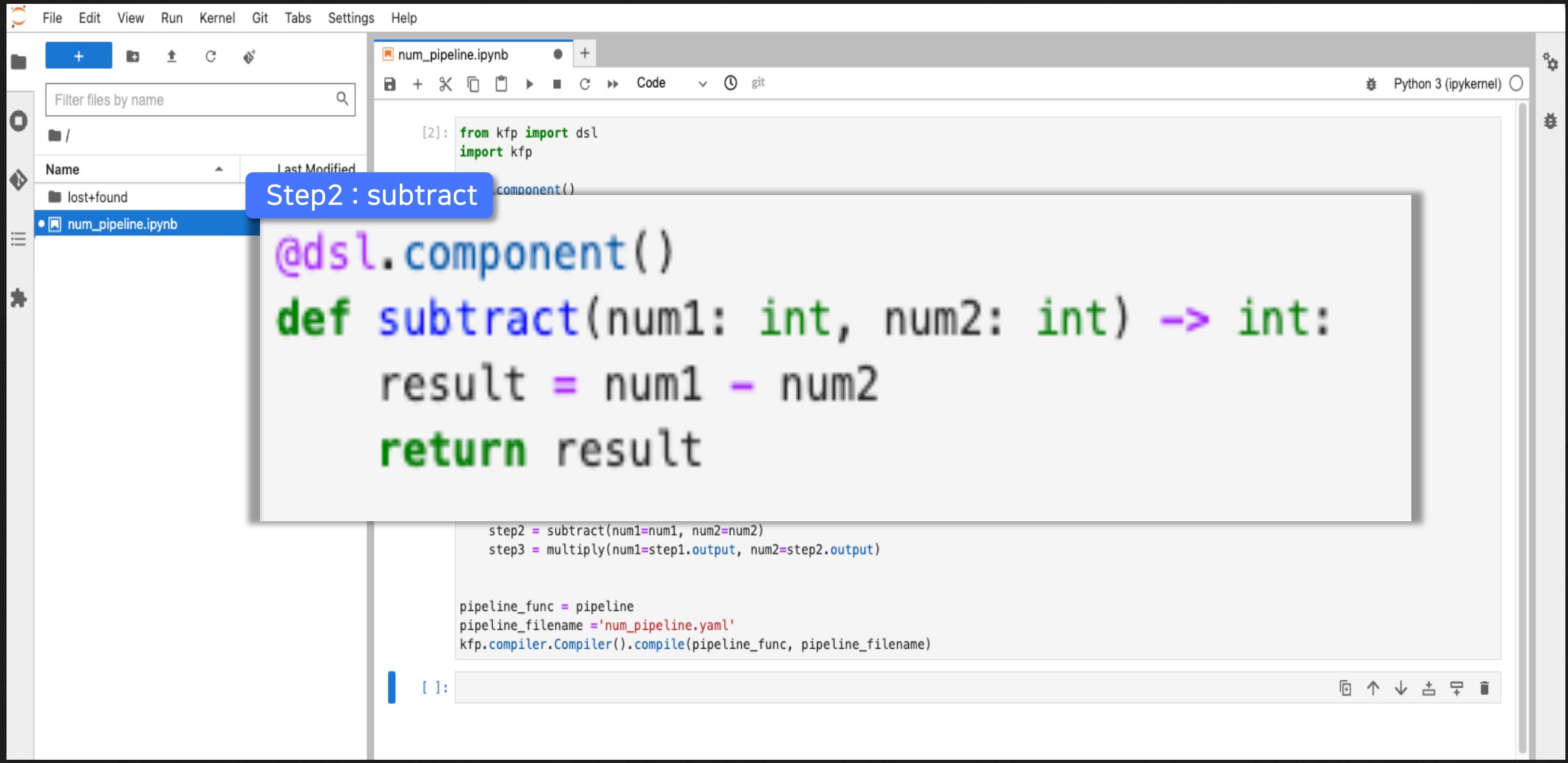
02 Kubeflow Components – Kubeflow Pipeline



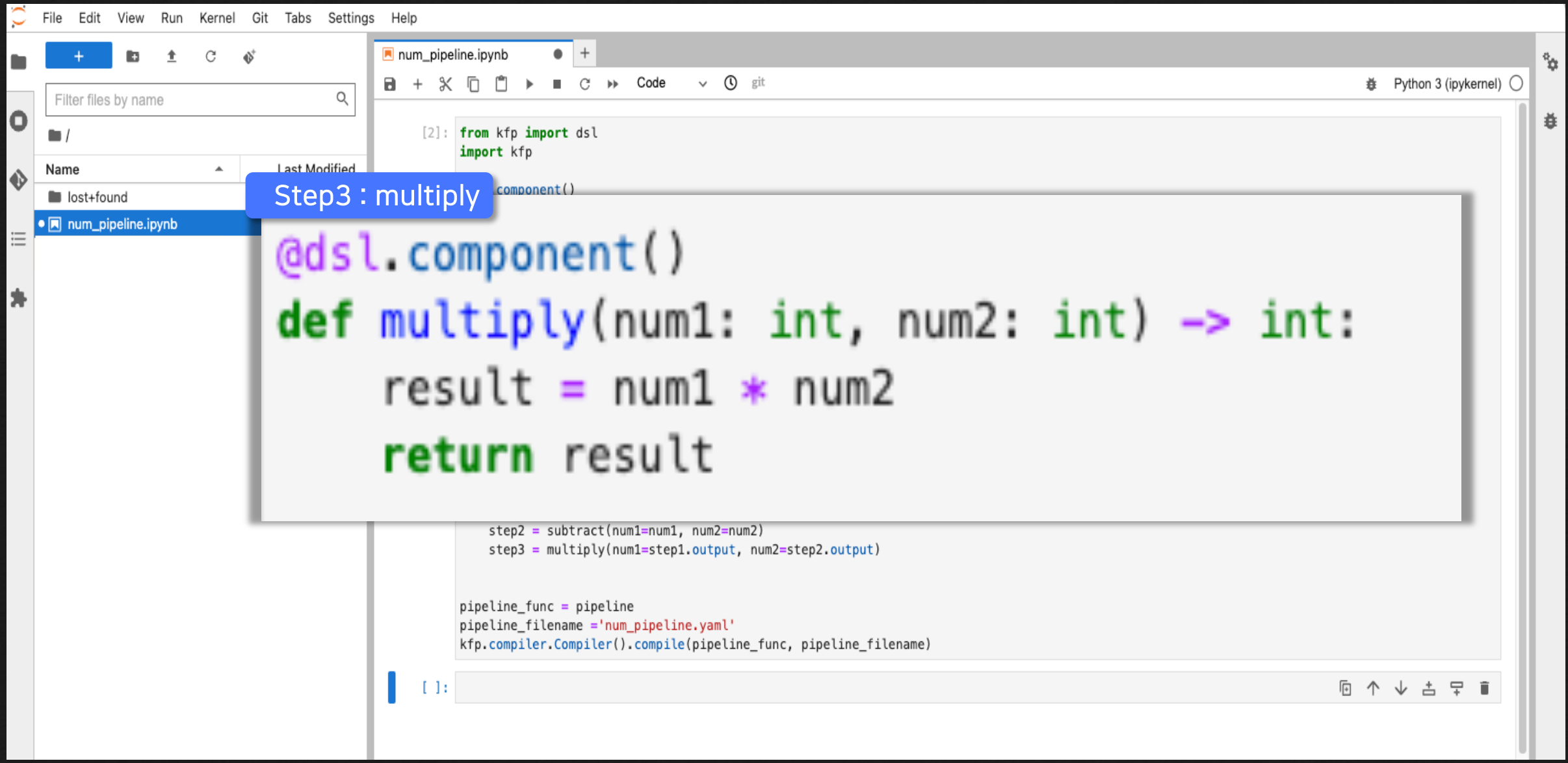
02 Kubeflow Components – Kubeflow Pipeline



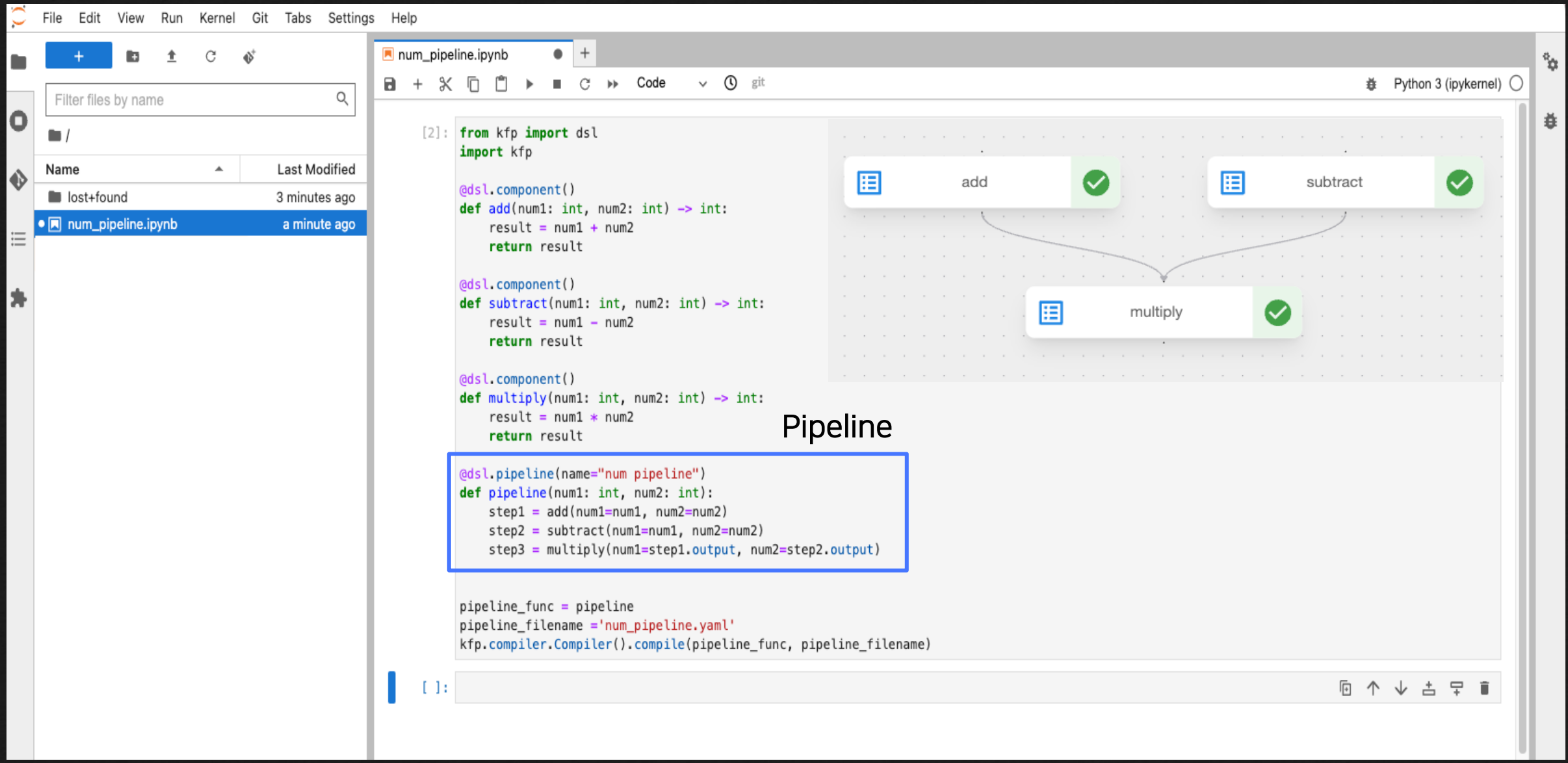
02 Kubeflow Components – Kubeflow Pipeline



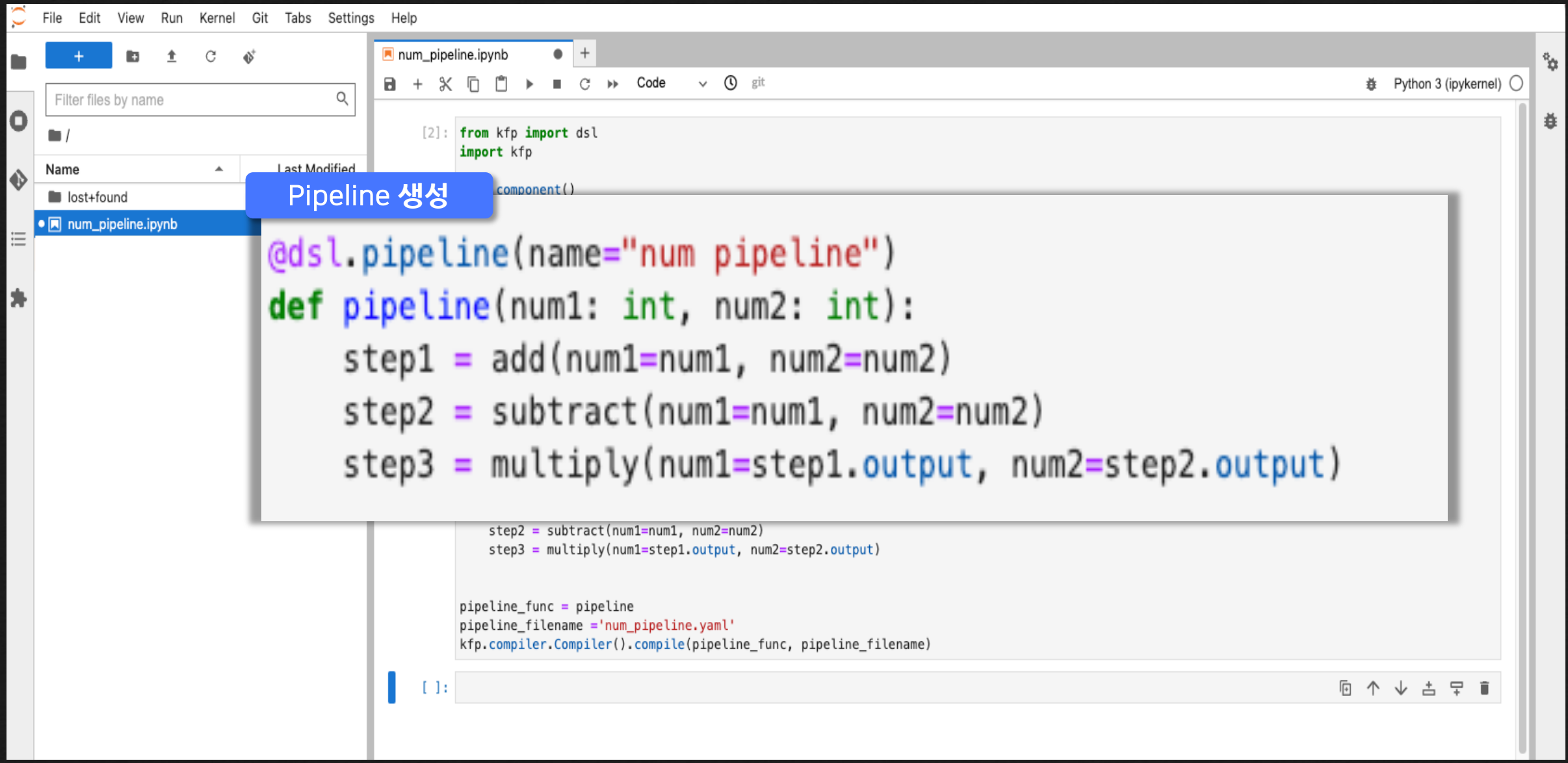
02 Kubeflow Components – Kubeflow Pipeline



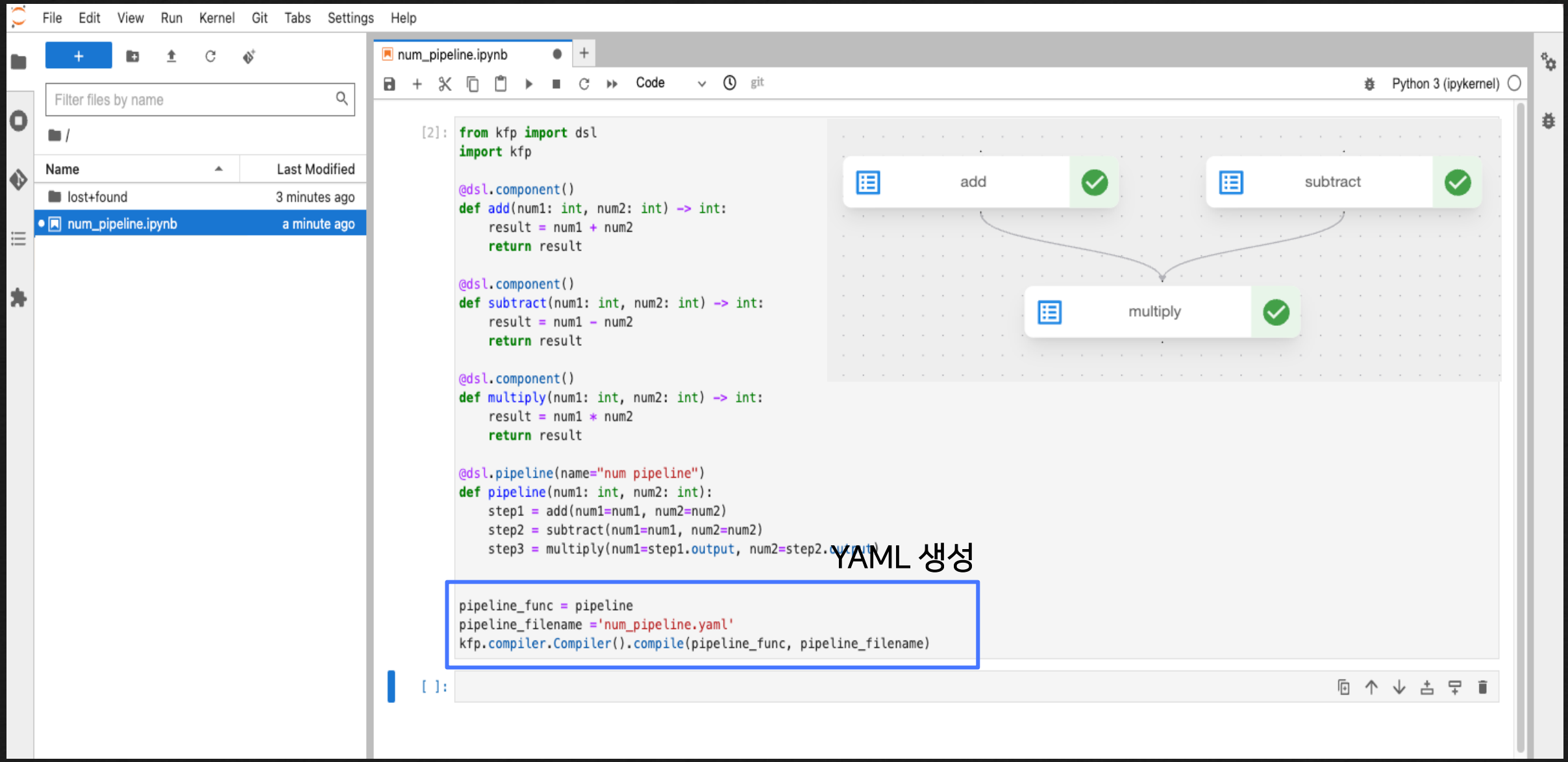
02 Kubeflow Components – Kubeflow Pipeline



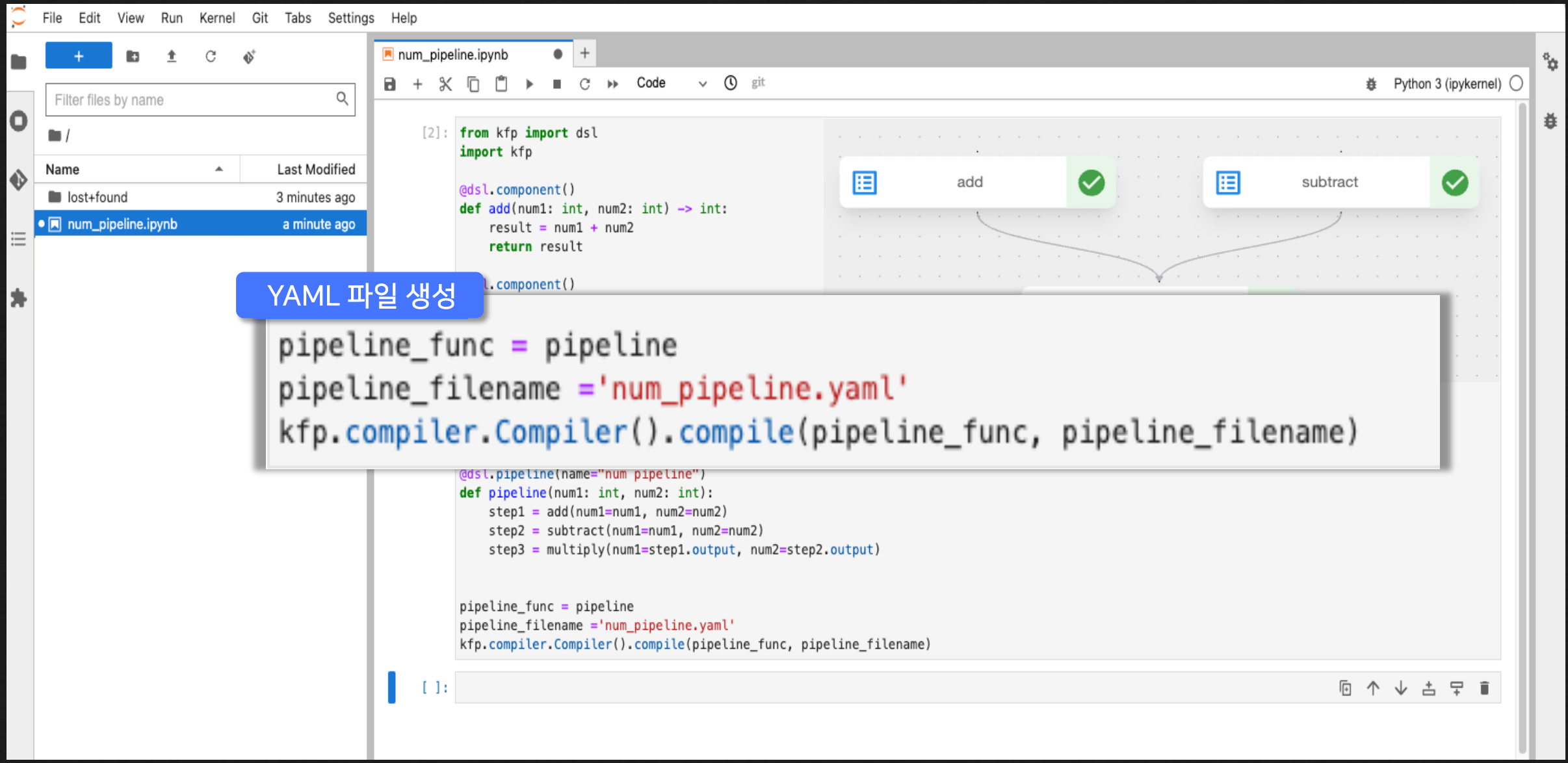
02 Kubeflow Components – Kubeflow Pipeline



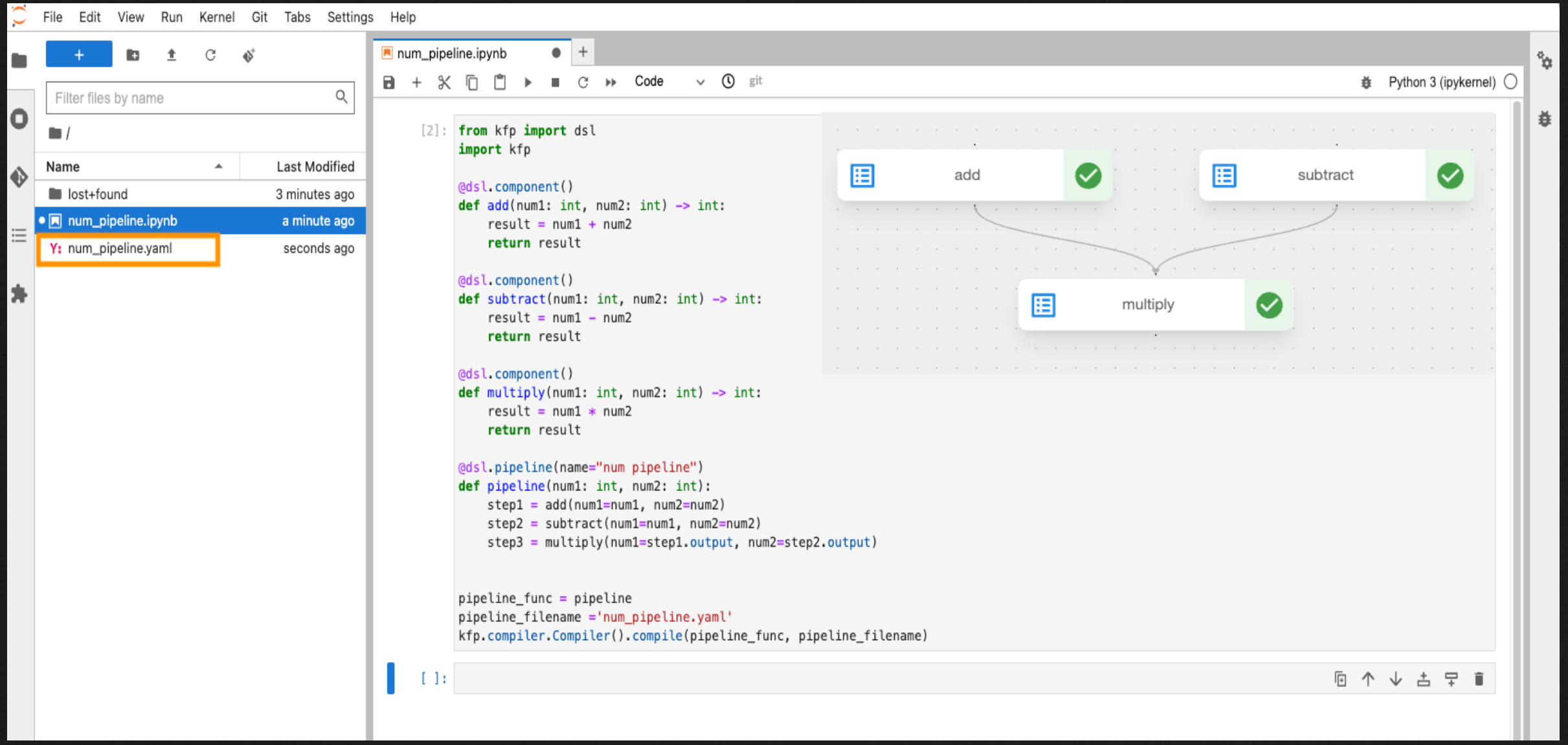
02 Kubeflow Components – Kubeflow Pipeline



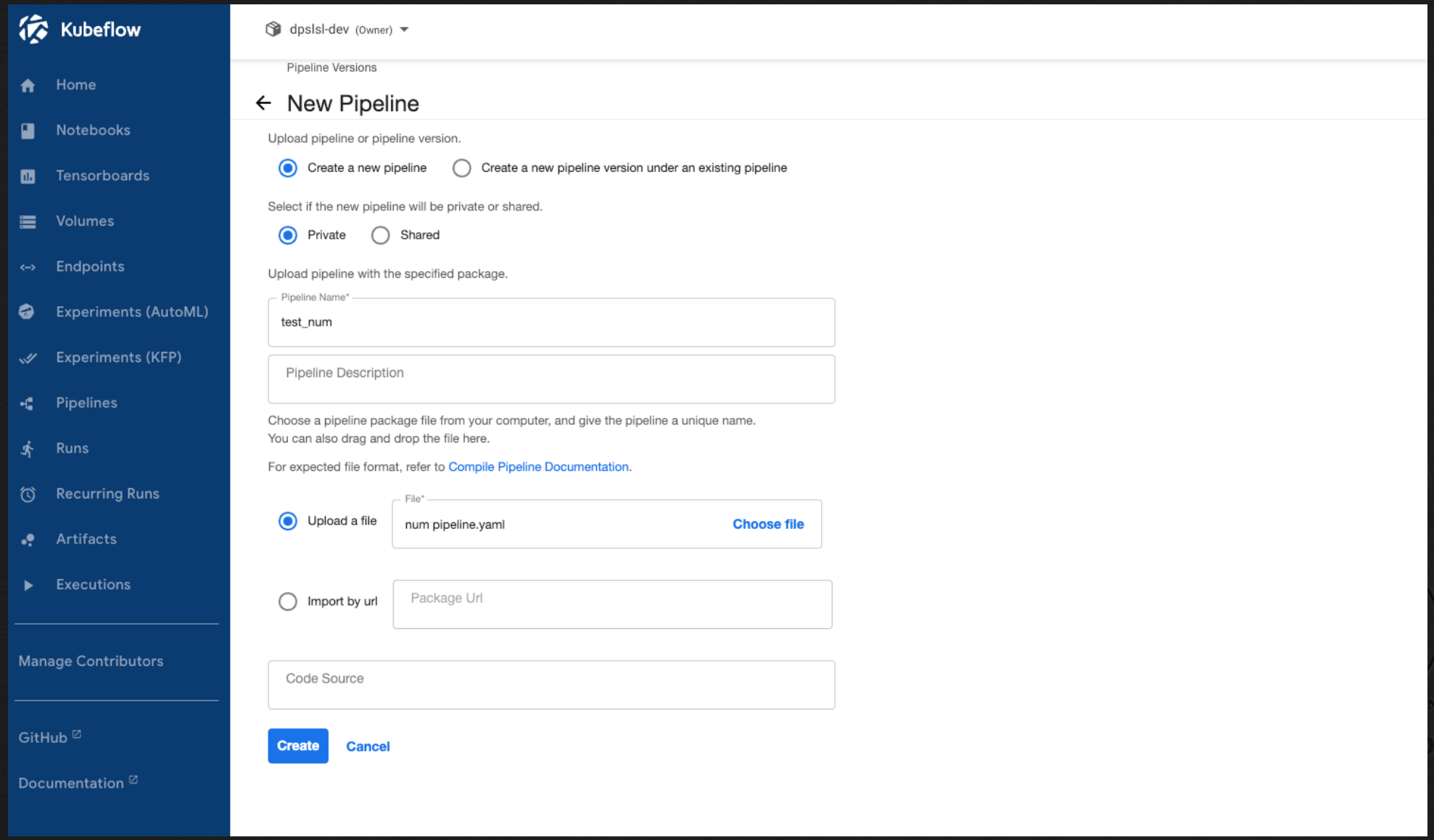
02 Kubeflow Components – Kubeflow Pipeline



02 Kubeflow Components – Kubeflow Pipeline

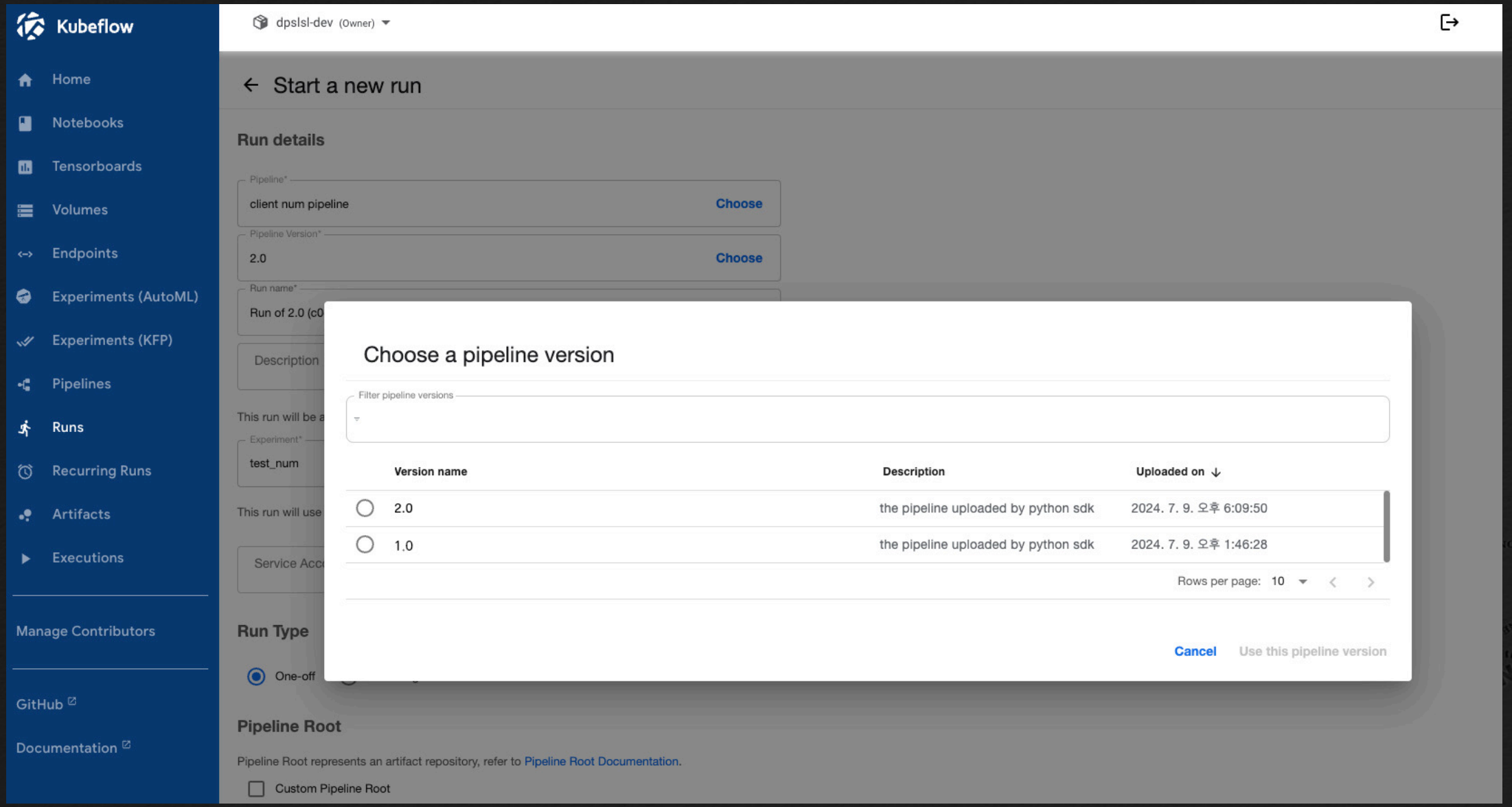


02 Kubeflow Components – Kubeflow Pipeline

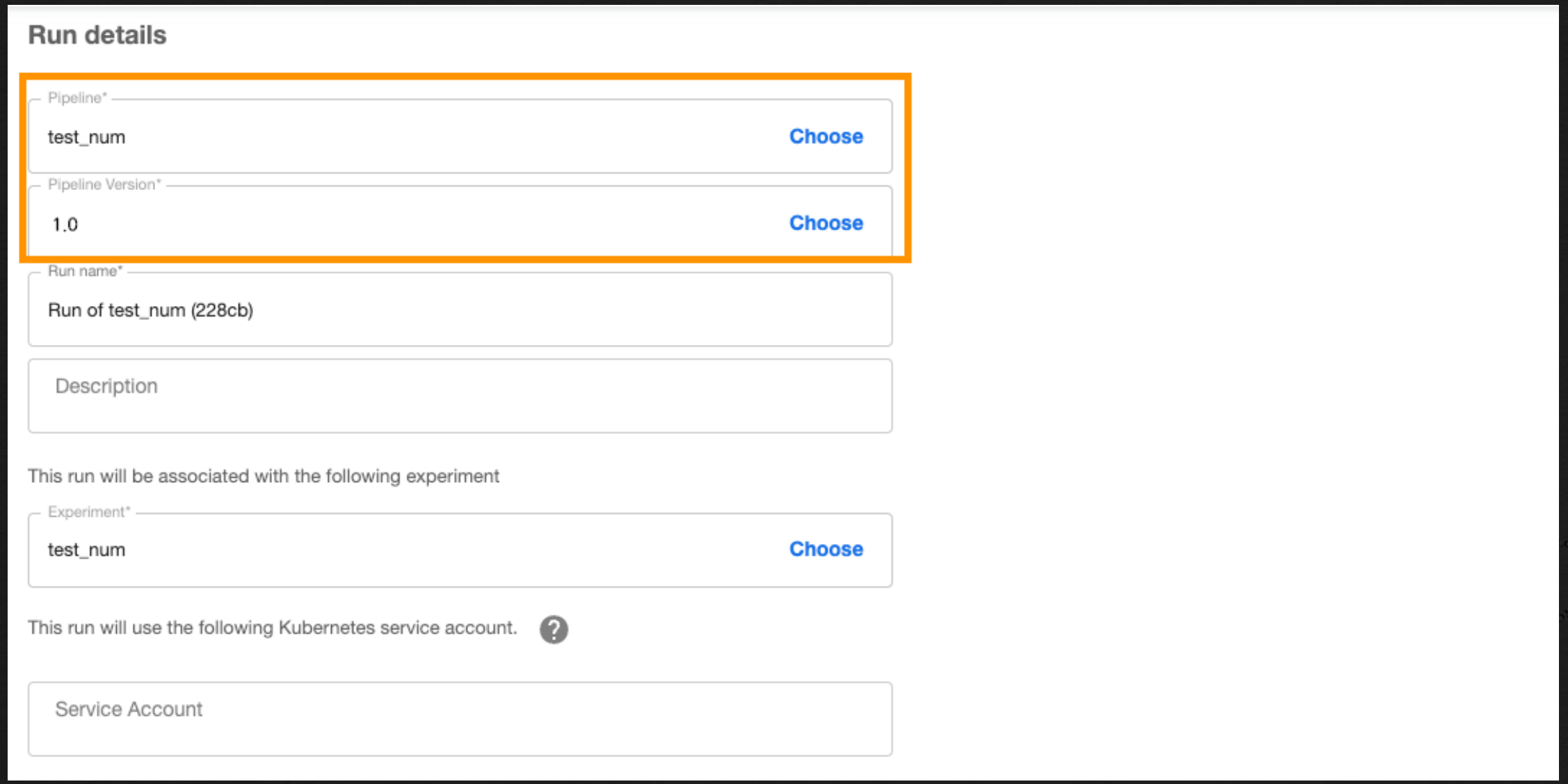


02 Kubeflow Components – Kubeflow Pipeline

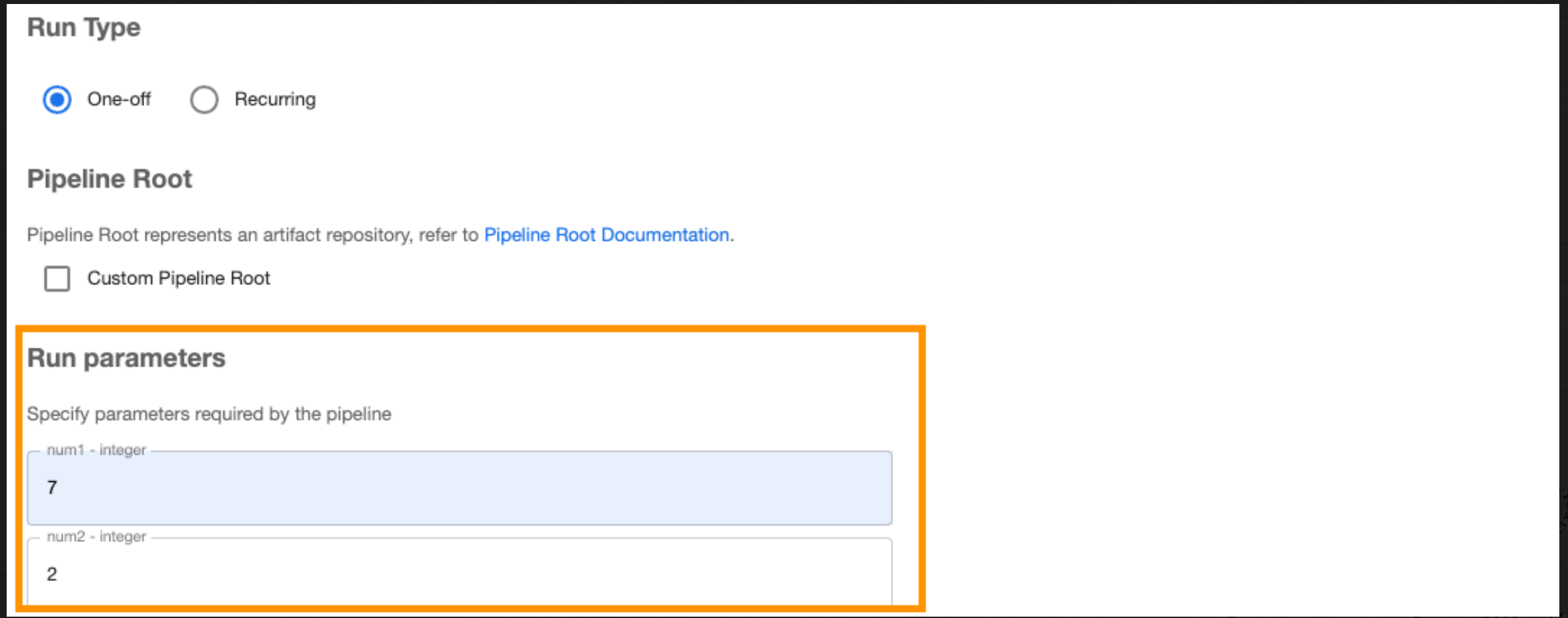
02 Kubeflow Components – Kubeflow Pipeline



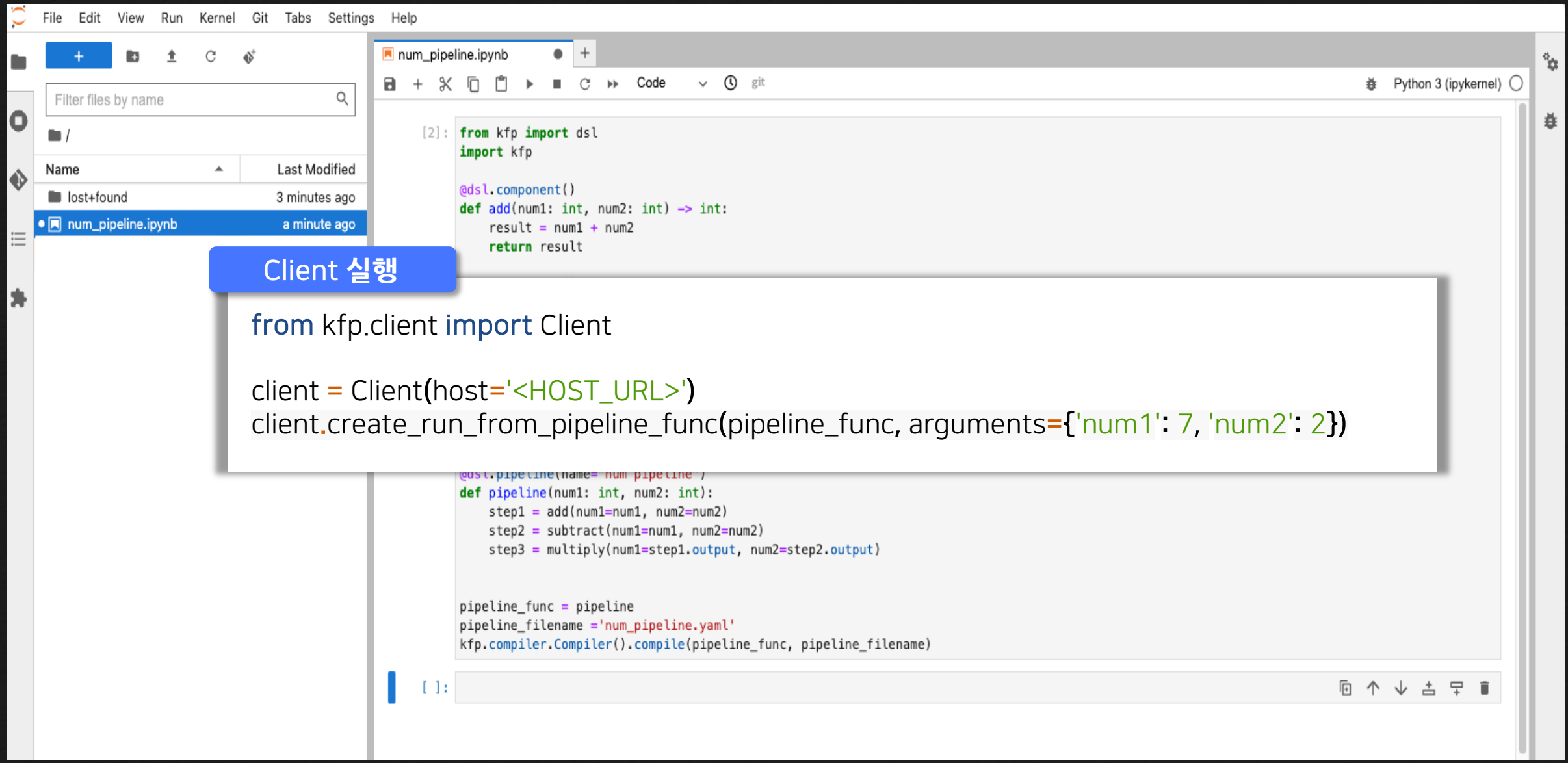
02 Kubeflow Components – Kubeflow Pipeline



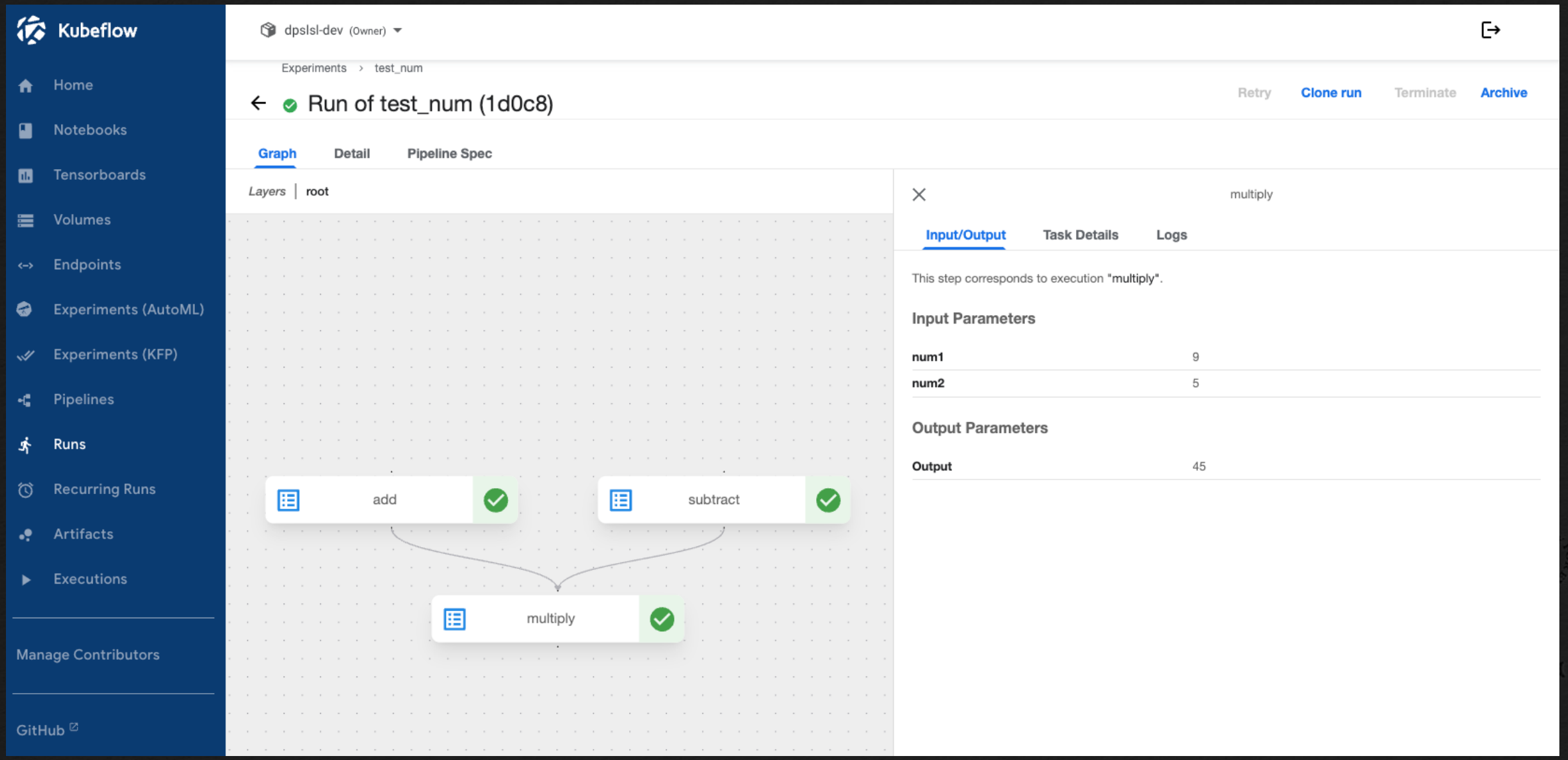
02 Kubeflow Components – Kubeflow Pipeline



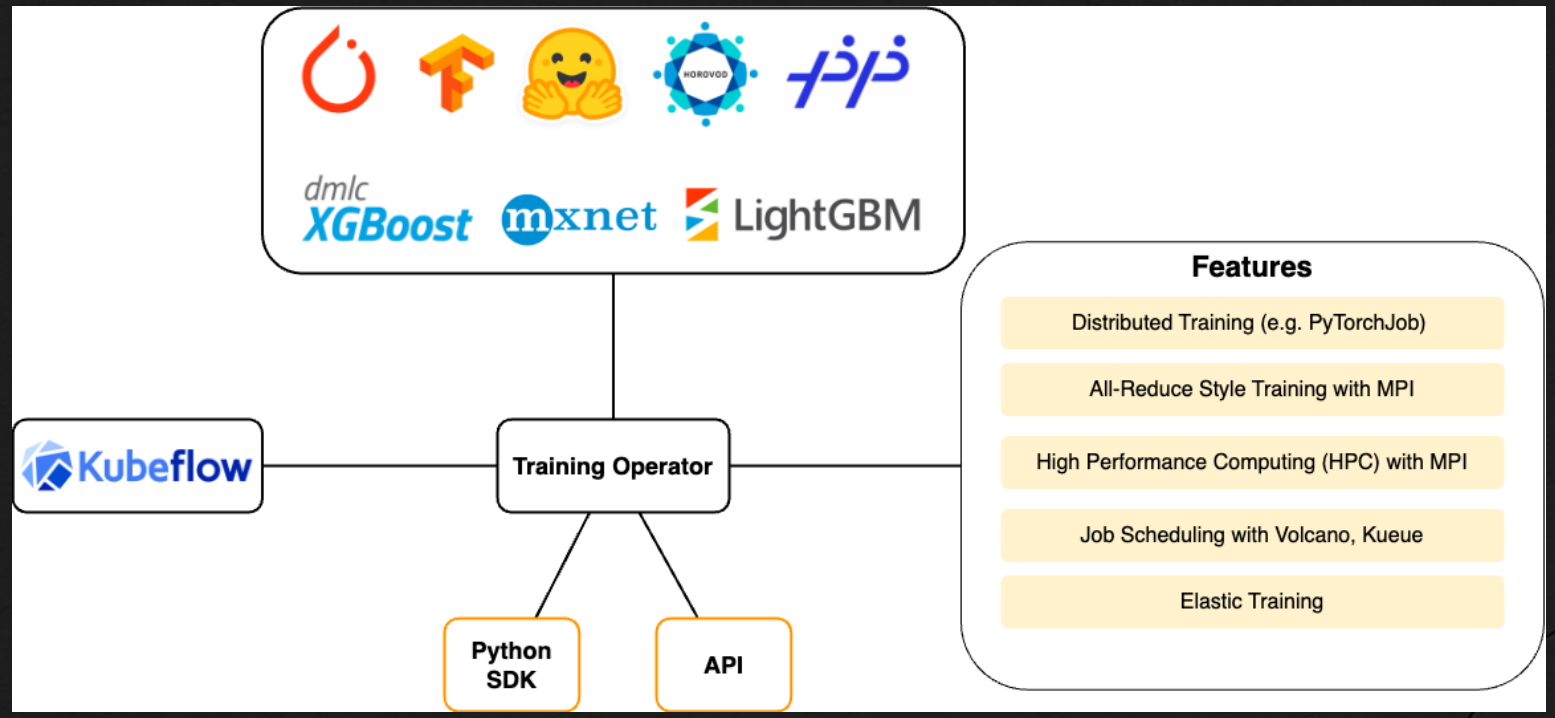
02 Kubeflow Components – Kubeflow Pipeline



02 Kubeflow Components – Kubeflow Pipeline



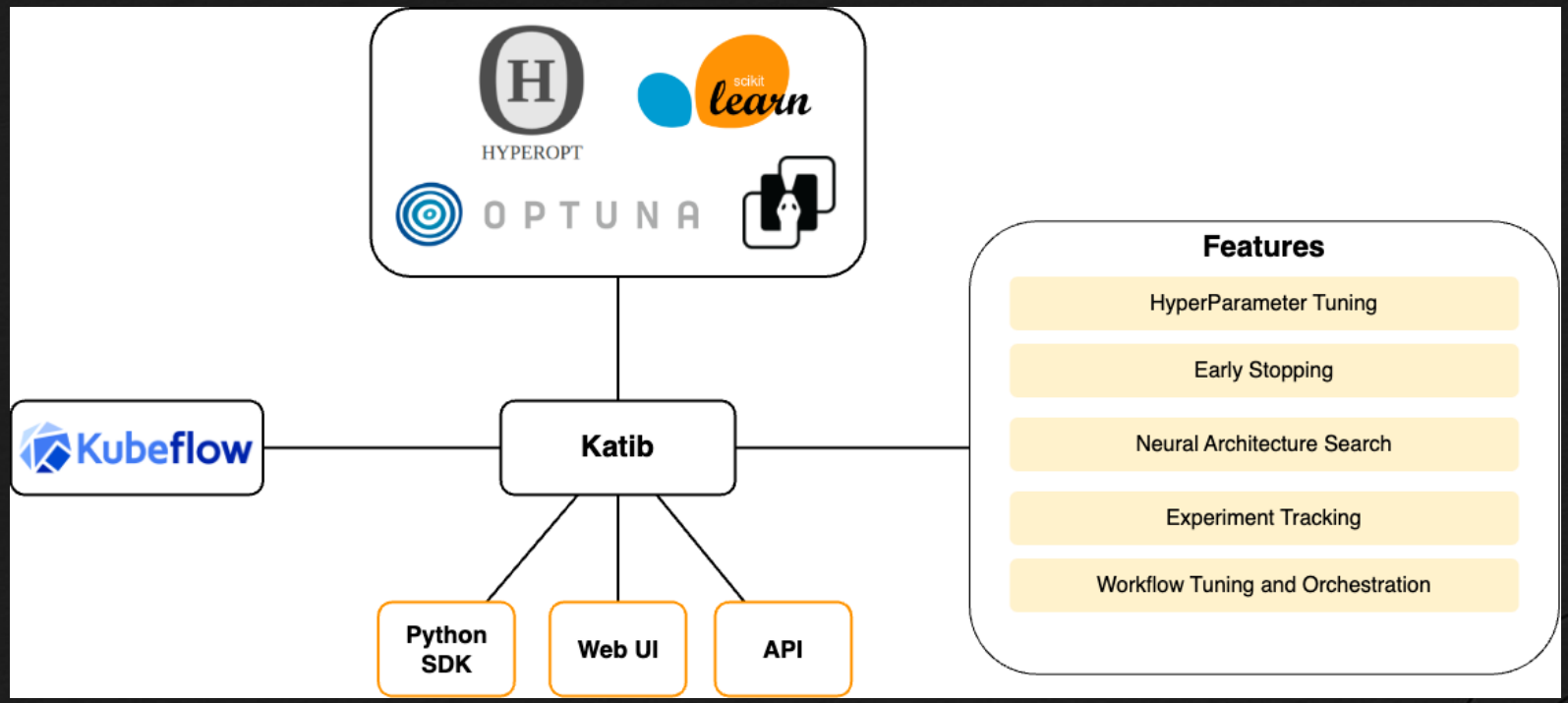
02 Kubeflow Components – Kubeflow Training Operator



- PyTorch, TensorFlow, XGBoost 등 다양한 ML 프레임워크를 사용하여 만든 ML 모델의 미세 조정 및 분산 학습을 위한 Kubernetes 기반 프로젝트
- HuggingFace , DeepSpeed와 같은 다른 ML 라이브러리를 Training Operator와 통합하여 Kubernetes에서 ML 학습 가능
- API와 Training Operator Python SDK를 이용하여 구성 가능

출처 : https://www.kubeflow.org/docs/components/training/overview/

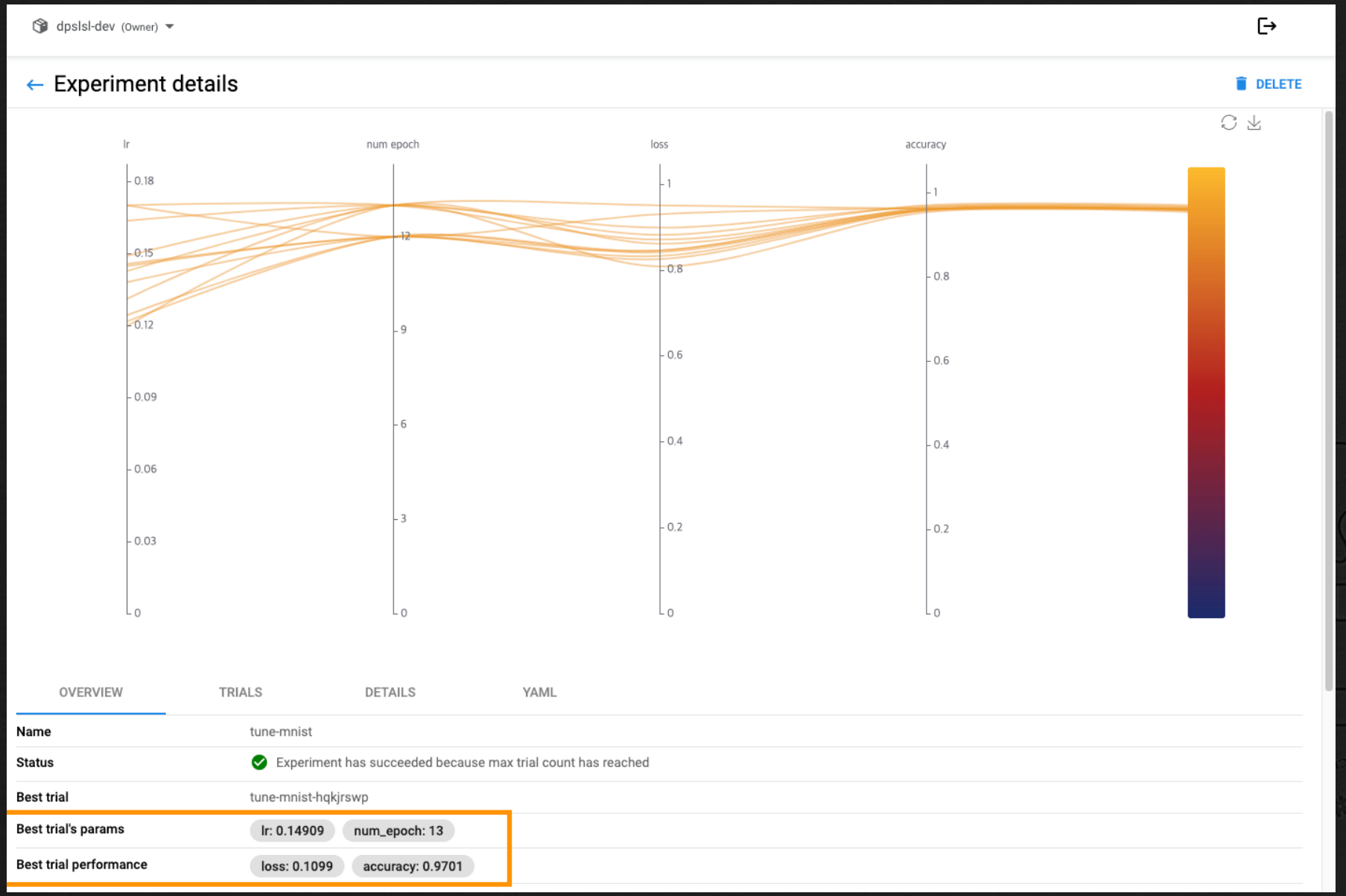
02 Kubeflow Components – Katib



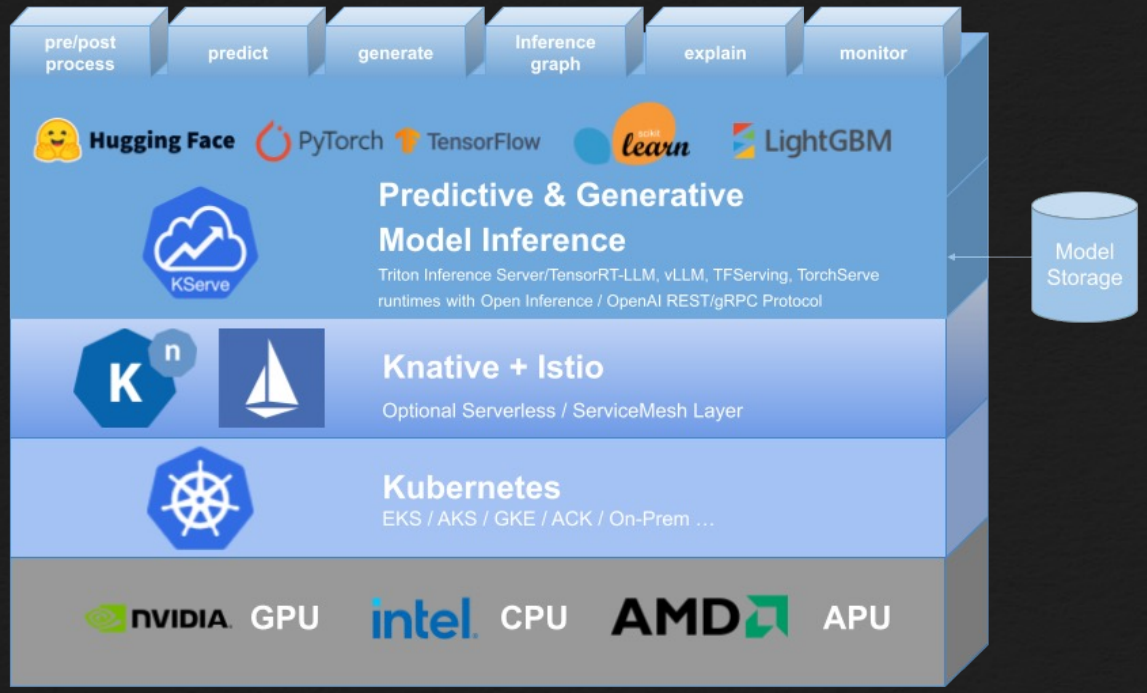
- AutoML을 위한 Kubernetes 기반 프로젝트
- Hyperparameter tuning, Early stopping 및 Neural Architecture Search(NAS)를 지원
- Bayesian optimization, Tree of Parzen Estimators, Random 등의 다양한 AutoML 알고리즘을 지원
- 최적화할 parameter, algorithm, early stopping point, metrix, maxTrialCount, maxFailedTrialCount 등을 통하여 조정 가능

출처 : https://www.kubeflow.org/docs/components/katib/overview/

02 Kubeflow Components – Katib



02 Kubeflow Components – KServe



KServe

- Scikit-Learn, XGBoost, Tensorflow, PyTorch와 같은 ML 프레임워크와 사용자 정의 모델을 사용하여 CPU, GPU에서 모델 추론이 가능한 서버리스 배포를 제공
- 자동 확장, 네트워킹 상태 검사, Scale to Zero 및 Canary Rollouts등의 기능을 ML 배포시에 제공
- Inference Service 스토리지 제공 : Google GCS, AWS S3, Azure Blob Storage, 로컬 컨테이너 파일시스템, PVC 형식

출처 : https://kserve.github.io/website/latest/

02 Kubeflow Components – KServe

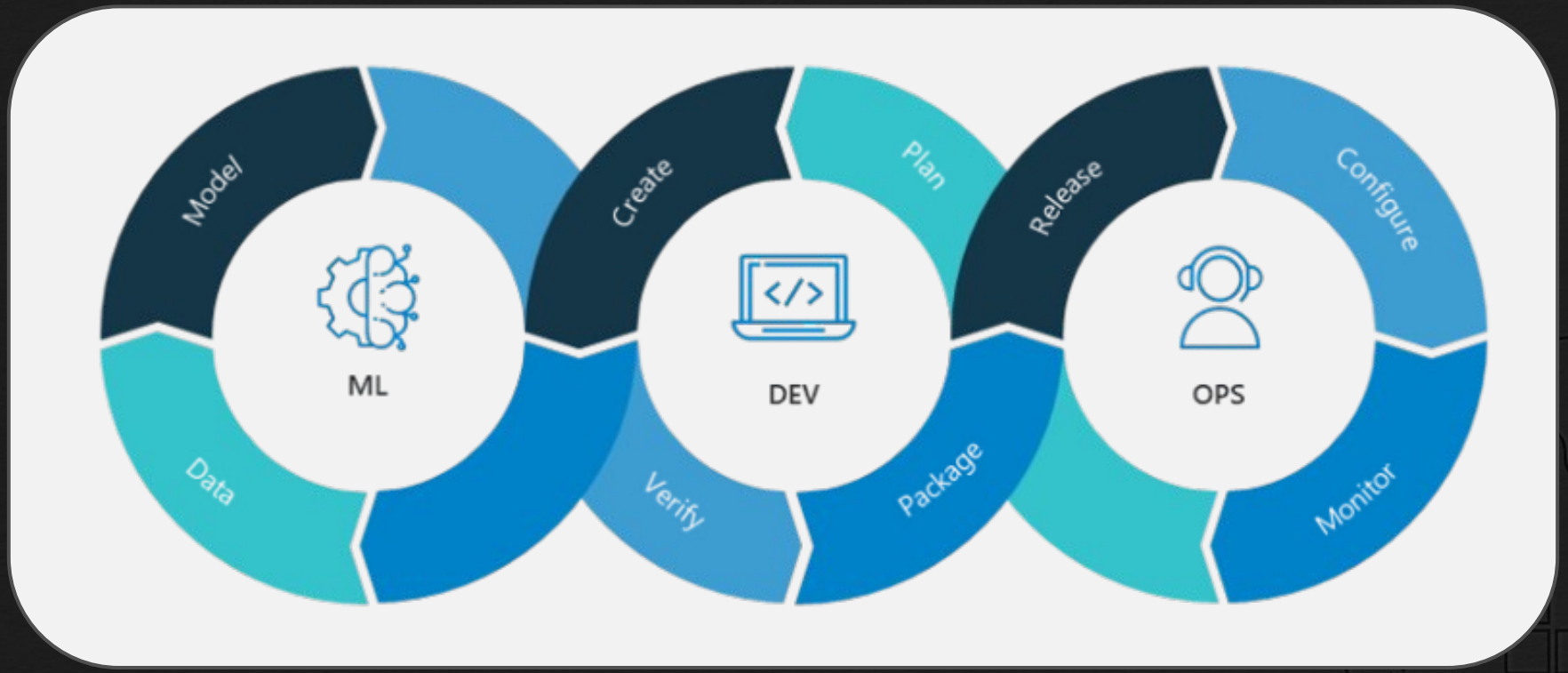
POST 요청

예측값 반환

요청	반환값
GET /v1/models	모델 목록
GET /v1/models/<model_name>	모델 상태
POST /v1/models/<model_name>:predict	예측
POST /v1/models/<model_name>:explain	예측 설명

03 MLOps

Machine Learning(ML) + Operations(Ops) = MLOps



출처 : https://blogs.nvidia.co.kr/blog/what-is-mlops/

03 MLOps

출처 : https://blogs.nvidia.co.kr/blog/what-is-mlops/

03 MLOps

DEVOCEAN

감사합니다.