Kubeflow Intro

Ala Raddaoui
Michal Jastrzebski
$whoami

- Ala Raddaoui
- Michal Jastrzebski
Current situation

- Data scientists must learn Devops
- Devops must learn machine learning
What is Kubeflow?

The Kubeflow project is dedicated to making deployments of machine learning (ML) workflows on Kubernetes simple, portable and scalable. Our goal is not to recreate other services, but to provide a straightforward way to deploy best-of-breed open-source systems for ML to diverse infrastructures. Anywhere you are running Kubernetes, you should be able to run Kubeflow.
So what is Kubeflow?

- Multiple components deployed together
  - JupyterHub
  - Tensorflow operator
  - PyTorch operator
  - Caffe2 operator
  - Katib
  - KVC
- Community focused on bringing ML to Kubernetes
Model lifecycle

- Setup infrastructure
- Develop model
- Train model
- Serve model
What do we deploy?

- Kubernetes 1.9.5
  - Deployed by Kubespray
- S3 over GCS
- Rook
  - Persistent volumes
- Kubeflow v. 0.1.0
Let’s go through an end to end example
Kubeflow installation
Model development
What do we train?

- **UNet**
  - Paper by Olaf Ronneberger, Philipp Fischer and Thomas Brox
  - Implementation by Tony Reina and Dina Suehiro, Intel AI

- **Dataset**
  - Courtesy of University of Pennsylvania BraTS team
Training
Without Kubeflow

Setup infrastructure
1. Deployment
2. Setup networking
3. Bootstrapping

Setup scheduling
1. Cluster spec
2. Which is which

Launch training
Launch training in each server
With Kubeflow

1. Create a docker image

2. Run training job
   1. Create ksonnet prototype
   2. Set params
   3. apply
Serving
Without Kubeflow

1. **Setup infrastructure**
   - 1. Deployment
   - 2. Setup networking
   - 3. Bootstrapping

2. **Deploy model**

3. **Setup load balancing**

4. **monitoring**
With Kubeflow

1. Instantiate ksonnet prototype
2. Set params
3. apply
What’s next

- Join Slack channel
- Join mailing list
- Subscribe to twitter account
- Look for Kubeflow talks
- Contribute!
Questions?