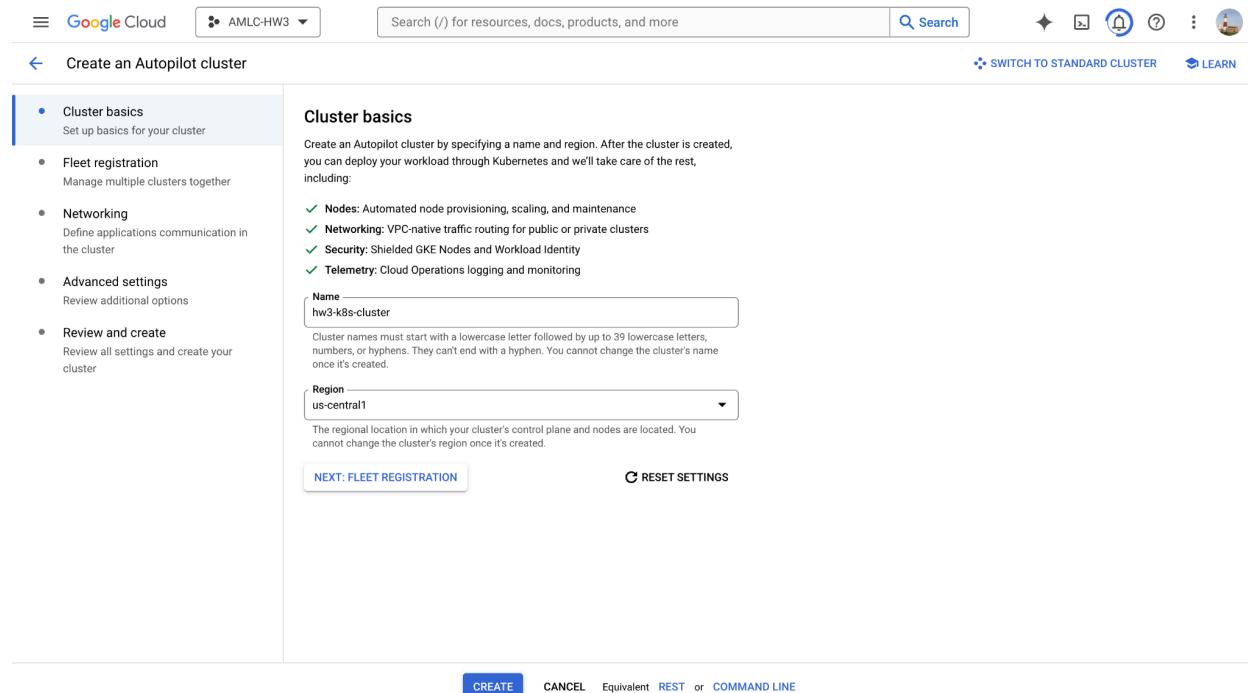
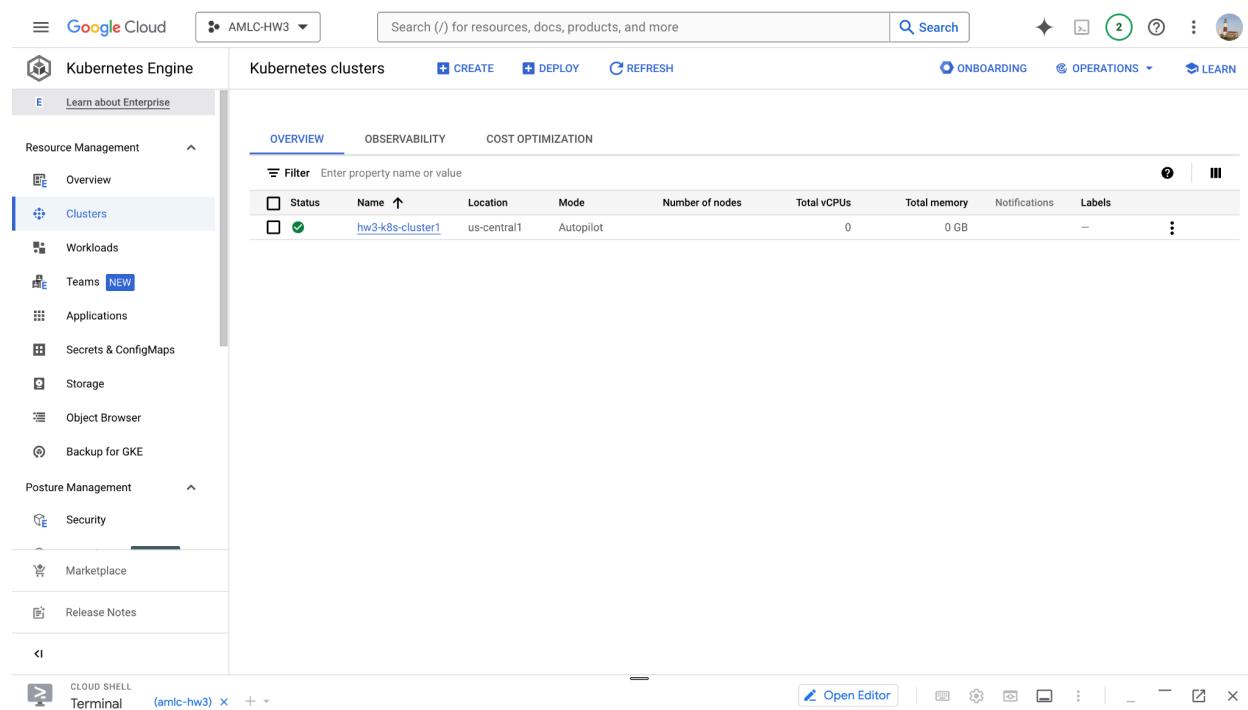


Samuel Braun
 Prof. Chung and Prof. Seelam
 Applied Machine Learning in the Cloud
 Spring 2024

1. Part 1, Steps 1-2



The screenshot shows the Google Cloud Platform interface for creating an Autopilot cluster. The left sidebar has a 'Cluster basics' section with options like 'Fleet registration', 'Networking', and 'Advanced settings'. The main panel shows the 'Cluster basics' configuration with a 'Name' field set to 'hw3-k8s-cluster' and a 'Region' dropdown set to 'us-central1'. Below the form are 'NEXT: FLEET REGISTRATION' and 'RESET SETTINGS' buttons, and at the bottom are 'CREATE', 'CANCEL', and command-line options.



The screenshot shows the Google Cloud Platform interface for the Kubernetes Engine. The left sidebar includes sections for Resource Management (Clusters, Workloads, Teams, Applications, Secrets & ConfigMaps, Storage, Object Browser, Backup for GKE), Posture Management (Security, Marketplace), and Release Notes. The main panel displays the 'Kubernetes clusters' table with one entry:

Status	Name	Location	Mode	Number of nodes	Total vCPUs	Total memory	Notifications	Labels
<input checked="" type="checkbox"/>	hw3-k8s-cluster1	us-central1	Autopilot	0	0 GB	—	⋮	

The screenshot shows the Google Cloud Kubernetes Engine interface. On the left, there's a sidebar with 'Resource Management' and 'Posture Management' sections. The 'Clusters' option under Resource Management is selected. A central modal window titled 'Connect to the cluster' provides instructions for connecting via command-line or dashboard. It includes a 'Command-line access' section with a terminal command and a 'RUN IN CLOUD SHELL' button, and a 'Cloud Console dashboard' section with a 'OPEN WORKLOADS DASHBOARD' button. At the bottom right of the modal is an 'OK' button.

2. Part 1, Step 2 - Part 2, Step 2

The screenshot shows a Cloud Shell terminal window. The session is named '(amlc-hw3)'. The terminal output shows the deployment of the Kubeflow training operator:

```
Welcome to Cloud Shell! Type "help" to get started.  
Your Cloud Platform project in this session is set to amlc-hw3.  
Use "gcloud config set project PROJECT_ID" to change to a different project.  
$ gcloud container clusters get-credentials hw3-k8s-cluster1 --region us-central1 --project amlc-hw3  
Fetching cluster endpoint and auth data  
kubeconfig entry generated for hw3-k8s-cluster1.  
$ kubectl apply -k "github.com/kubeflow/training-operator/manifests/standalone?ref=v1.7.0"  
namespace/kubeflow created  
customresourcedefinition.apirextensions.k8s.io/mpijobs.kubeflow.org created  
customresourcedefinition.apirextensions.k8s.io/mxjobs.kubeflow.org created  
customresourcedefinition.apirextensions.k8s.io/paddlejobs.kubeflow.org created  
customresourcedefinition.apirextensions.k8s.io/pytorchjobs.kubeflow.org created  
customresourcedefinition.apirextensions.k8s.io/tfjobs.kubeflow.org created  
customresourcedefinition.apirextensions.k8s.io/xgbboostjobs.kubeflow.org created  
serviceaccount/training-operator created  
clusterrole/rbac.authorization.k8s.io/training-operator created  
clusterrolebinding.rbac.authorization.k8s.io/training-operator created  
service/training-operator created  
deployment.apps/training-operator created  
$ kubectl config set-context --current --namespace=kubeflow  
Context "gke_amlc-hw3_us-central1_hw3-k8s-cluster1" modified.  
$ cd [REDACTED]
```

3. Part 2, Step 3

4. Part 2, Step 4

The screenshot shows the Google Cloud Platform (GCP) interface for the Kubernetes Engine. The left sidebar is collapsed, showing various management sections like Resource Management, Storage, and Posture Management. The main content area is titled 'Storage' and displays 'PERSISTENT VOLUME CLAIMS'. A table lists one persistent volume claim: 'model-pvc' (Name), 'Bound' (Phase), 'pvc-8cb0fcf1-9dc3-4cc6-a8c6-525c9099a899' (Volume), 'standard' (Storage class), 'kubeflow' (Namespace), and 'hw3-k8s-cluster' (Cluster). The top navigation bar includes a search bar, a 'Search' button, and several icons for navigation and help.

Name	Phase	Volume	Storage class	Namespace	Cluster
model-pvc	Bound	pvc-8cb0fcf1-9dc3-4cc6-a8c6-525c9099a899	standard	kubeflow	hw3-k8s-cluster

Google Cloud AMLC-HW3 Search (/) for resources, docs, products, and more

Kubernetes Engine Persistent volume claim details REFRESH EDIT DELETE KUBECTL LEARN

Learn about Enterprise

All Fleets

Resource Management

- Overview
- Clusters
- Workloads
- Teams NEW
- Applications
- Secrets & ConfigMaps
- Storage
- Object Browser
- Backup for GKE

Posture Management

- Marketplace
- Release Notes

CLOUD SHELL Terminal (amlc-hw3) Open Editor

model-pvc

DETAILS EVENTS YAML

Cluster hw3-k8s-cluster

Namespace kubeflow

Created Apr 30, 2024, 6:03:22 AM

Labels No labels set

Annotations

- pv.kubernetes.io/bind-completed: yes
- pv.kubernetes.io/bound-by-controller: yes
- volume.beta.kubernetes.io/storage-provisioner: pd.csi.storage.gke.io
- volume.kubernetes.io/storage-provisioner: pd.csi.storage.gke.io

Phase Bound

Label selector No labels set

Requested access modes Read Write Once

Storage class standard

Resources

- Storage requested 10Gi, capacity 10Gi
- Volume pvc-8cb0fcf1-9dc3-4cc6-a8c6-525c9099a899

SHOW ALL ANNOTATIONS

5. Part 2, Step 5

Google Cloud AMLC-HW3 Search (/) for resources, docs, products, and more

CLOUD SHELL Terminal (amlc-hw3) Open Editor

```
slb2250@cloudshell:~/amlc/hw3/training (amlc-hw3)$ clear
slb2250@cloudshell:~/amlc/hw3/training (amlc-hw3)$ kubectl get pods
NAME          READY   STATUS    RESTARTS   AGE
pytorch-seelam-master-0   0/1     Pending   0          54s
pytorch-seelam-worker-0   0/1     Pending   0          54s
training-operator-66d795fc75-t8r4s  1/1     Running   0          10h
slb2250@cloudshell:~/amlc/hw3/training (amlc-hw3)$
```

Google Cloud AMLC-HW3 Search (/) for resources, docs, products, and more

CLOUD SHELL Terminal (amlc-hw3) +

```
training-operator-66b795fc75-t8r4s 1/1 Running 0 10h
slb2250@cloudshell:~/amlc/hw3/training (amlc-hw3)$ kubectl get pods
NAME READY STATUS RESTARTS AGE
pytorch-seelam-master-0 0/1 ContainerCreating 0 4m9s
pytorch-seelam-worker-0 0/1 Init:0/0 1 0 4m9s
training-operator-66b795fc75-t8r4s 1/1 Running 0 10h
slb2250@cloudshell:~/amlc/hw3/training (amlc-hw3)$ kubectl get pods
NAME READY STATUS RESTARTS AGE
pytorch-seelam-master-0 0/1 ContainerCreating 0 4m34s
pytorch-seelam-worker-0 0/1 Init:0/0 1 0 4m34s
training-operator-66b795fc75-t8r4s 1/1 Running 0 10h
slb2250@cloudshell:~/amlc/hw3/training (amlc-hw3)$ kubectl get pods
NAME READY STATUS RESTARTS AGE
pytorch-seelam-master-0 0/1 ContainerCreating 0 4m49s
pytorch-seelam-worker-0 0/1 Init:0/0 1 0 4m49s
training-operator-66b795fc75-t8r4s 1/1 Running 0 10h
slb2250@cloudshell:~/amlc/hw3/training (amlc-hw3)$ kubectl get pods
NAME READY STATUS RESTARTS AGE
pytorch-seelam-master-0 1/1 Running 0 5m32s
pytorch-seelam-worker-0 1/1 Running 0 5m32s
training-operator-66b795fc75-t8r4 1/1 Running 0 10h
slb2250@cloudshell:~/amlc/hw3/training (amlc-hw3)$ kubectl log pytorch-seelam-master-0
Download http://yann.lecun.com/exdb/mnist/train-labels-idx3-ubyte.gz
Download http://yann.lecun.com/exdb/mnist/train-images-idx3-ubyte.gz
Download http://yann.lecun.com/exdb/mnist/train-labels-idx3-ubyte.gz to ..../data/MNIST/raw/train-labels-idx3-ubyte.gz
100%[=====] 9912422/9912422 [0:00<0:00, 67173527.29it/s]
Extracting ..../data/MNIST/raw/train-labels-idx3-ubyte.gz to ..../data/MNIST/raw

Download http://yann.lecun.com/exdb/mnist/train-labels-idx1-ubyte.gz
Download http://yann.lecun.com/exdb/mnist/train-labels-idx1-ubyte.gz to ..../data/MNIST/raw/train-labels-idx1-ubyte.gz
100%[=====] 28881/28881 [0:00<0:00, 15828523.95it/s]
Extracting ..../data/MNIST/raw/train-labels-idx1-ubyte.gz to ..../data/MNIST/raw

Downloading http://yann.lecun.com/exdb/mnist/t10k-labels-idx3-ubyte.gz
Downloading http://yann.lecun.com/exdb/mnist/t10k-labels-idx3-ubyte.gz to ..../data/MNIST/raw/t10k-labels-idx3-ubyte.gz
100%[=====] 1648877/1648877 [0:00<0:00, 13663530.08it/s]
Extracting ..../data/MNIST/raw/t10k-labels-idx3-ubyte.gz to ..../data/MNIST/raw

Downloading http://yann.lecun.com/exdb/mnist/t10k-labels-idx1-ubyte.gz
Downloading http://yann.lecun.com/exdb/mnist/t10k-labels-idx1-ubyte.gz to ..../data/MNIST/raw/t10k-labels-idx1-ubyte.gz
100%[=====] 4542/4542 [0:00<0:00, 6155259.70it/s]
Extracting ..../data/MNIST/raw/t10k-labels-idx1-ubyte.gz to ..../data/MNIST/raw

Train Epoch: 1 [0/60000 (0%)] Loss: 2.305400
Train Epoch: 1 [640/60000 (1%)] Loss: 1.359780
Train Epoch: 1 [1280/60000 (2%)] Loss: 0.830670
Train Epoch: 1 [1920/60000 (3%)] Loss: 0.605961
slb2250@cloudshell:~/amlc/hw3/training (amlc-hw3)$
```

Google Cloud AMLC-HW3 Search (/) for resources, docs, products, and more

CLOUD SHELL Terminal (amlc-hw3) + ×

```
Train Epoch: 1 [5760/60000 (10%)] Loss: 0.208878
Train Epoch: 1 [16000/60000 (11%)] Loss: 0.279327
slb2250@cloudshell:~/amlc/hw3$ python3 training.py --model=leap --pycharm=selam-master=0
downloading http://yann.lecun.com/exdb/mnist/train-images-idx3-ubyte.gz
downloading http://yann.lecun.com/exdb/mnist/train-images-idx3-ubyte.gz to ./data/MNIST/raw/train-images-idx3-ubyte.gz
100% [██████████] 9912422/9912422 [00:00<0:00, 67173527.29it/s]
Extracting ./data/MNIST/raw/train-images-idx3-ubyte.gz to ./data/MNIST/raw

downloading http://yann.lecun.com/exdb/mnist/train-labels-idx1-ubyte.gz
downloading http://yann.lecun.com/exdb/mnist/train-labels-idx1-ubyte.gz to ./data/MNIST/raw/train-labels-idx1-ubyte.gz
100% [██████████] 28881/28881 [00:00<0:00, 1582853.95it/s]
Extracting ./data/MNIST/raw/train-labels-idx1-ubyte.gz to ./data/MNIST/raw

downloading http://yann.lecun.com/exdb/mnist/t10k-images-idx3-ubyte.gz
downloading http://yann.lecun.com/exdb/mnist/t10k-images-idx3-ubyte.gz to ./data/MNIST/raw/t10k-images-idx3-ubyte.gz
100% [██████████] 1648877/1648877 [00:00<0:00, 13663530.08it/s]
Extracting ./data/MNIST/raw/t10k-images-idx3-ubyte.gz to ./data/MNIST/raw

downloading http://yann.lecun.com/exdb/mnist/t10k-labels-idx1-ubyte.gz
downloading http://yann.lecun.com/exdb/mnist/t10k-labels-idx1-ubyte.gz to ./data/MNIST/raw/t10k-labels-idx1-ubyte.gz
100% [██████████] 4542/4542 [00:00<0:00, 6155259.70it/s]
Extracting ./data/MNIST/raw/t10k-labels-idx1-ubyte.gz to ./data/MNIST/raw

Train Epoch: 1 [0/60000 (0%)] Loss: 2.305400
Train Epoch: 1 [640/60000 (1%)] Loss: 1.359780
Train Epoch: 1 [1280/60000 (2%)] Loss: 0.830670
Train Epoch: 1 [1920/60000 (3%)] Loss: 0.605961
Train Epoch: 1 [2560/60000 (4%)] Loss: 0.345934
Train Epoch: 1 [3200/60000 (5%)] Loss: 0.446331
Train Epoch: 1 [3840/60000 (6%)] Loss: 0.306768
Train Epoch: 1 [4480/60000 (7%)] Loss: 0.279325
Train Epoch: 1 [5120/60000 (8%)] Loss: 0.555025
Train Epoch: 1 [5760/60000 (9%)] Loss: 0.279326
Train Epoch: 1 [6400/60000 (10%)] Loss: 0.279327
Train Epoch: 1 [7040/60000 (12%)] Loss: 0.327207
Train Epoch: 1 [7680/60000 (13%)] Loss: 0.204888
Train Epoch: 1 [8320/60000 (14%)] Loss: 0.220855
Train Epoch: 1 [8960/60000 (15%)] Loss: 0.273643
Train Epoch: 1 [9600/60000 (16%)] Loss: 0.097318
Train Epoch: 1 [10240/60000 (17%)] Loss: 0.248318
Train Epoch: 1 [10880/60000 (18%)] Loss: 0.112893
Train Epoch: 1 [11520/60000 (19%)] Loss: 0.439983
Train Epoch: 1 [12160/60000 (20%)] Loss: 0.244582
Train Epoch: 1 [12800/60000 (21%)] Loss: 0.157299
Train Epoch: 1 [13440/60000 (22%)] Loss: 0.221483
Train Epoch: 1 [14080/60000 (23%)] Loss: 0.157299
Train Epoch: 1 [14720/60000 (25%)] Loss: 0.418896
Train Epoch: 1 [15360/60000 (26%)] Loss: 0.168725
Train Epoch: 1 [16000/60000 (27%)] Loss: 0.110782
Train Epoch: 1 [16640/60000 (28%)] Loss: 0.185906
Train Epoch: 1 [17280/60000 (29%)] Loss: 0.054832
Train Epoch: 1 [17920/60000 (30%)] Loss: 0.167925
Train Epoch: 1 [18560/60000 (31%)] Loss: 0.195231
Train Epoch: 1 [19200/60000 (32%)] Loss: 0.282548
Train Epoch: 1 [19840/60000 (33%)] Loss: 0.042161
Train Epoch: 1 [20480/60000 (34%)] Loss: 0.042161
Train Epoch: 1 [21120/60000 (35%)] Loss: 0.217422
Train Epoch: 1 [21760/60000 (36%)] Loss: 0.004475
Train Epoch: 1 [22400/60000 (37%)] Loss: 0.066934
Train Epoch: 1 [23040/60000 (38%)] Loss: 0.208920
Train Epoch: 1 [23680/60000 (39%)] Loss: 0.229430
Train Epoch: 1 [24320/60000 (41%)] Loss: 0.014528
Train Epoch: 1 [24960/60000 (42%)] Loss: 0.130294
Train Epoch: 1 [25600/60000 (43%)] Loss: 0.071659
Train Epoch: 1 [26240/60000 (44%)] Loss: 0.177676
Train Epoch: 1 [26880/60000 (45%)] Loss: 0.383814
Train Epoch: 1 [27520/60000 (46%)] Loss: 0.263688
Train Epoch: 1 [28160/60000 (47%)] Loss: 0.123490
Train Epoch: 1 [28800/60000 (48%)] Loss: 0.099680
Train Epoch: 1 [29440/60000 (49%)] Loss: 0.037860
Train Epoch: 1 [30080/60000 (50%)] Loss: 0.149224
Train Epoch: 1 [30720/60000 (51%)] Loss: 0.042093
Train Epoch: 1 [31360/60000 (52%)] Loss: 0.105343
Train Epoch: 1 [32000/60000 (53%)] Loss: 0.232382
Train Epoch: 1 [32640/60000 (54%)] Loss: 0.049167
Train Epoch: 1 [33280/60000 (55%)] Loss: 0.007577
Train Epoch: 1 [33920/60000 (57%)] Loss: 0.021158
Train Epoch: 1 [34560/60000 (58%)] Loss: 0.023109
Train Epoch: 1 [35200/60000 (59%)] Loss: 0.225845
Train Epoch: 1 [35840/60000 (60%)] Loss: 0.185864
Train Epoch: 1 [36480/60000 (61%)] Loss: 0.061325
Train Epoch: 1 [37120/60000 (62%)] Loss: 0.071801
Train Epoch: 1 [37760/60000 (63%)] Loss: 0.162966
Train Epoch: 1 [38400/60000 (64%)] Loss: 0.135675
Train Epoch: 1 [39040/60000 (65%)] Loss: 0.072343
slb2250@cloudshell:~/amlc/hw3$
```

Google Cloud AMLC-HW3 Search (/) for resources, docs, products, and more

CLOUD SHELL Terminal (amlc-hw3) × (amlc-hw3) × +

```
Train Epoch: 1 [10240/60000 (17%)] Loss: 0.248318
Train Epoch: 1 [10880/60000 (18%)] Loss: 0.112893
Train Epoch: 1 [11520/60000 (19%)] Loss: 0.439983
Train Epoch: 1 [12160/60000 (20%)] Loss: 0.244582
Train Epoch: 1 [12800/60000 (21%)] Loss: 0.157299
Train Epoch: 1 [13440/60000 (22%)] Loss: 0.221483
Train Epoch: 1 [14080/60000 (23%)] Loss: 0.157299
Train Epoch: 1 [14720/60000 (25%)] Loss: 0.418896
Train Epoch: 1 [15360/60000 (26%)] Loss: 0.168725
Train Epoch: 1 [16000/60000 (27%)] Loss: 0.110782
Train Epoch: 1 [16640/60000 (28%)] Loss: 0.185906
Train Epoch: 1 [17280/60000 (29%)] Loss: 0.054832
Train Epoch: 1 [17920/60000 (30%)] Loss: 0.167925
Train Epoch: 1 [18560/60000 (31%)] Loss: 0.195231
Train Epoch: 1 [19200/60000 (32%)] Loss: 0.282548
Train Epoch: 1 [19840/60000 (33%)] Loss: 0.042161
Train Epoch: 1 [20480/60000 (34%)] Loss: 0.042161
Train Epoch: 1 [21120/60000 (35%)] Loss: 0.217422
Train Epoch: 1 [21760/60000 (36%)] Loss: 0.004475
Train Epoch: 1 [22400/60000 (37%)] Loss: 0.066934
Train Epoch: 1 [23040/60000 (38%)] Loss: 0.208920
Train Epoch: 1 [23680/60000 (39%)] Loss: 0.229430
Train Epoch: 1 [24320/60000 (41%)] Loss: 0.014528
Train Epoch: 1 [24960/60000 (42%)] Loss: 0.130294
Train Epoch: 1 [25600/60000 (43%)] Loss: 0.071659
Train Epoch: 1 [26240/60000 (44%)] Loss: 0.177676
Train Epoch: 1 [26880/60000 (45%)] Loss: 0.383814
Train Epoch: 1 [27520/60000 (46%)] Loss: 0.263688
Train Epoch: 1 [28160/60000 (47%)] Loss: 0.123490
Train Epoch: 1 [28800/60000 (48%)] Loss: 0.099680
Train Epoch: 1 [29440/60000 (49%)] Loss: 0.037860
Train Epoch: 1 [30080/60000 (50%)] Loss: 0.149224
Train Epoch: 1 [30720/60000 (51%)] Loss: 0.042093
Train Epoch: 1 [31360/60000 (52%)] Loss: 0.105343
Train Epoch: 1 [32000/60000 (53%)] Loss: 0.232382
Train Epoch: 1 [32640/60000 (54%)] Loss: 0.049167
Train Epoch: 1 [33280/60000 (55%)] Loss: 0.007577
Train Epoch: 1 [33920/60000 (57%)] Loss: 0.021158
Train Epoch: 1 [34560/60000 (58%)] Loss: 0.023109
Train Epoch: 1 [35200/60000 (59%)] Loss: 0.225845
Train Epoch: 1 [35840/60000 (60%)] Loss: 0.185864
Train Epoch: 1 [36480/60000 (61%)] Loss: 0.061325
Train Epoch: 1 [37120/60000 (62%)] Loss: 0.071801
Train Epoch: 1 [37760/60000 (63%)] Loss: 0.162966
Train Epoch: 1 [38400/60000 (64%)] Loss: 0.135675
Train Epoch: 1 [39040/60000 (65%)] Loss: 0.072343
slb2250@cloudshell:~/amlc/hw3$
```

Google Cloud AMLC-HW3 Search 3

CLOUD SHELL Terminal (amlc-hw3) (amlc-hw3) +

Open Editor

Train Epoch: 1 [32640/60000 (54%)] Loss: 0.135167
Train Epoch: 1 [33280/60000 (55%)] Loss: 0.067577
Train Epoch: 1 [33920/60000 (57%)] Loss: 0.021159
Train Epoch: 1 [34560/60000 (58%)] Loss: 0.023109
Train Epoch: 1 [35200/60000 (59%)] Loss: 0.225845
Train Epoch: 1 [35840/60000 (60%)] Loss: 0.185864
Train Epoch: 1 [36480/60000 (61%)] Loss: 0.061325
Train Epoch: 1 [37120/60000 (62%)] Loss: 0.071801
Train Epoch: 1 [37760/60000 (63%)] Loss: 0.162966
Train Epoch: 1 [38400/60000 (64%)] Loss: 0.135675
Train Epoch: 1 [39040/60000 (65%)] Loss: 0.072343
Train Epoch: 1 [39680/60000 (66%)] Loss: 0.041530
Train Epoch: 1 [40320/60000 (67%)] Loss: 0.076386
Train Epoch: 1 [40960/60000 (68%)] Loss: 0.120102
Train Epoch: 1 [41600/60000 (69%)] Loss: 0.126077
Train Epoch: 1 [42240/60000 (70%)] Loss: 0.101517
Train Epoch: 1 [42880/60000 (71%)] Loss: 0.091000
Train Epoch: 1 [43520/60000 (72%)] Loss: 0.209517
Train Epoch: 1 [44160/60000 (74%)] Loss: 0.058008
Train Epoch: 1 [44800/60000 (75%)] Loss: 0.128887
Train Epoch: 1 [45440/60000 (76%)] Loss: 0.179389
Train Epoch: 1 [46080/60000 (77%)] Loss: 0.041446
Train Epoch: 1 [46720/60000 (78%)] Loss: 0.151364
Train Epoch: 1 [47360/60000 (79%)] Loss: 0.129901
Train Epoch: 1 [48000/60000 (80%)] Loss: 0.058679
Train Epoch: 1 [48640/60000 (81%)] Loss: 0.019105
Train Epoch: 1 [49280/60000 (82%)] Loss: 0.041951
Train Epoch: 1 [49920/60000 (83%)] Loss: 0.080363
Train Epoch: 1 [50560/60000 (84%)] Loss: 0.075957
Train Epoch: 1 [51200/60000 (85%)] Loss: 0.307334
Train Epoch: 1 [51840/60000 (86%)] Loss: 0.015220
Train Epoch: 1 [52480/60000 (87%)] Loss: 0.044241
Train Epoch: 1 [53120/60000 (88%)] Loss: 0.125356
Train Epoch: 1 [53760/60000 (90%)] Loss: 0.077792
Train Epoch: 1 [54400/60000 (91%)] Loss: 0.090643
Train Epoch: 1 [55040/60000 (92%)] Loss: 0.022916
Train Epoch: 1 [55680/60000 (93%)] Loss: 0.121202
Train Epoch: 1 [56320/60000 (94%)] Loss: 0.089443
Train Epoch: 1 [56960/60000 (95%)] Loss: 0.079569
Train Epoch: 1 [57600/60000 (96%)] Loss: 0.139670
Train Epoch: 1 [58240/60000 (97%)] Loss: 0.021290
Train Epoch: 1 [58880/60000 (98%)] Loss: 0.027833
Train Epoch: 1 [59520/60000 (99%)] Loss: 0.001188

Test set: Average loss: 0.0466, Accuracy: 9844/10000 (98%)

mlb2250@cloudshell:~/amlc/hw3/training (amlc-hw3)\$

6. Part 3, Steps 6-8

Google Cloud AMLC-HW3 Search (/) for resources, docs, products, and more

CLOUD SHELL Terminal (amlc-hw3) Open Editor

```
cloud2250@cloudshell:~/amlc/hw3/inference (amlc-hw3)$ docker build -t gcr.io/mlc-hw3/mnist-inference:latest .
[+] Building 0.6s (10/10) FINISHED
   Digest: sha256:2f911e2866173a52104dc16b5e42b7069c2eba05eb78556d18b1ca665d0dc445
   Layer 0: FROM docker.io/library/python:3.8-slim
   Layer 1: >-- transferring dockerfile: 203B
   Layer 2: >-- [internal] load metadata for docker.io/library/python:3.8-slim
   Layer 3: >-- [internal] load .dockerignore
   Layer 4: >-- [internal] transferring context: 2B
   Layer 5: >-- [internal] load build context
   Layer 6: >-- [internal] transferring context: 1.37KB
   Layer 7: CACHE(2/5) WORKDIR /app
   Layer 8: CACHE(3/5) COPY requirements.txt /app
   Layer 9: CACHE(4/5) RUN pip install --no-cache-dir -r requirements.txt
   Layer 10: COPY .
   Layer 11: EXPAND
   Layer 12: EXPORTING TO IMAGE
   Layer 13: EXPORTING LAYERS
   Layer 14: >>> writing image sha256:dbd64290ad90e769abc84f6b56e418d8d2a408ce6475acf516b9b39e1
   Layer 15: >>> naming to gcr.io/mlc/hw3/mnist-inference:latest
   Layer 16: push refers to repository [gcr.io/mlc-hw3/mnist-inference]
759c632a289: Pushed
d7d1a8dd977: Layer already exists
8a39e580d00: Layer already exists
28c500545f: Layer already exists
b6011f4711434: Layer already exists
8d8e7f754ef8: Layer already exists
cb0bec46633: Layer already exists
7a75d57a5024: Layer already exists
52ec5a4316fa: Layer already exists
latest: digest: sha256:46e538c307bb6172f533add0bf1c81ee3b4f3edd643c0bf3f673ad09a051a363 size: 2204
cloud2250@cloudshell:~/amlc/hw3/inference (amlc-hw3)$ kubectl create -f infer.yaml
deployment.apps/mnist-inference created
service/mnist-inference created
cloud2250@cloudshell:~/amlc/hw3/inference (amlc-hw3)$ kubectl get pods
NAME          READY   STATUS    RESTARTS   AGE
mnist-inference-6cffdd7f7-xtfsc   0/1   Pending   0          7s
training-operator-66b795fc7-dsmc2  1/1   Running   0          139m
cloud2250@cloudshell:~/amlc/hw3/inference (amlc-hw3)$ kubectl get pods
NAME          READY   STATUS    RESTARTS   AGE
mnist-inference-6cffdd7f7-xtfsc   0/1   Pending   0          31s
training-operator-66b795fc7-dsmc2  1/1   Running   0          139m
cloud2250@cloudshell:~/amlc/hw3/inference (amlc-hw3)$ kubectl get pods
NAME          READY   STATUS    RESTARTS   AGE
mnist-inference-6cffdd7f7-xtfsc   0/1   Pending   0          99s
training-operator-66b795fc7-dsmc2  1/1   Running   0          140m
cloud2250@cloudshell:~/amlc/hw3/inference (amlc-hw3)$ kubectl get pods
NAME          READY   STATUS    RESTARTS   AGE
cloud2250@cloudshell:~/amlc/hw3/inference (amlc-hw3)$
```

Google Cloud AMLC-HW3 Search (/) for resources, docs, products, and more

Kubernetes Engine Workloads Refresh Deploy Create Job New Delete Operations Learn

All Fleets

Resource Management Overview Observability Cost Optimization

Workloads Filter Is system object: False

Name	Status	Type	Pods	Namespace	Cluster
mnist-inference	OK	Deployment	1/1	kubeflow	hw3-k8s-cluster
training-operator	OK	Deployment	1/1	kubeflow	hw3-k8s-cluster

Posture Management Marketplace Release Notes

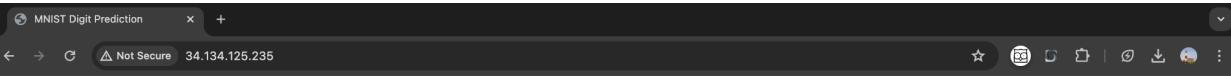
CLOUD SHELL Terminal (amlc-hw3) Open Editor

CLOUD SHELL Terminal (amlc-hw3) + ×

```

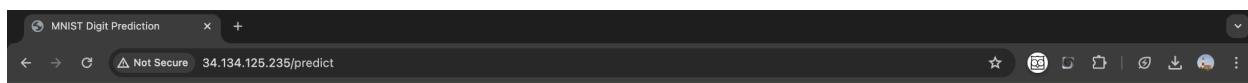
--> [5/5] COPY . /app
--> exporting to image
-->   EXPANDING layers
-->     >> 0 layers from sha256:dbd1e64290ad90e769abc84f6b56e418d8d56d2a408ce6475acfac518b9b39e1
-->     >> naming to gcr.io/amlc-hw3/mnist-inference:latest
579ce32a289: Pushed
The push refers to repository [gcr.io/amlc-hw3/mnist-inference]
579ce32a289: Pushed
9d1e1a8dd77: Layer already exists
8a3419e580fe: Layer already exists
28be0e023bd5f: Layer already exists
b60a1f4754ef: Layer already exists
8a3419e580fe: Layer already exists
bb08ec46633: Layer already exists
7a75d57a5024: Layer already exists
52ec5a4316fa: Layer already exists
latest: digest: sha256:46e538c307bb6172f533add0bf1c8lee3b4f3edd643c0bf3f673ad09a051a363 size: 2204
slb2250@cloudshell:~/amlc-hw3$ kubectl create -f infer.yaml
deployment.apps/mnist-inference created
service/mnist-inference created
slb2250@cloudshell:~/amlc-hw3$ kubectl get pods
NAME          READY   STATUS    RESTARTS   AGE
mnist-inference-6cffdd7ff7-xtfec  0/1   Pending   0          7s
training-operator-66b795fc75-dsmc2  1/1   Pending   0          10m
slb2250@cloudshell:~/amlc-hw3$ kubectl get pods
NAME          READY   STATUS    RESTARTS   AGE
mnist-inference-6cffdd7ff7-xtfec  0/1   Pending   0          31s
training-operator-66b795fc75-dsmc2  1/1   Running   0          139m
slb2250@cloudshell:~/amlc-hw3$ kubectl get pods
NAME          READY   STATUS    RESTARTS   AGE
mnist-inference-6cffdd7ff7-xtfec  0/1   Pending   0          99s
training-operator-66b795fc75-dsmc2  1/1   Running   0          140m
slb2250@cloudshell:~/amlc-hw3$ kubectl get pods
NAME          READY   STATUS    RESTARTS   AGE
mnist-inference-6cffdd7ff7-xtfec  0/1   ContainerCreating   0          3m26s
training-operator-66b795fc75-dsmc2  1/1   Running   0          142m
slb2250@cloudshell:~/amlc-hw3$ kubectl get pods
NAME          READY   STATUS    RESTARTS   AGE
mnist-inference-6cffdd7ff7-xtfec  1/1   Running   0          8m46s
training-operator-66b795fc75-dsmc2  1/1   Running   0          147m
slb2250@cloudshell:~/amlc-hw3$ kubectl get svc
NAME           TYPE      CLUSTER-IP      EXTERNAL-IP      PORT(S)        AGE
mnist-inference   LoadBalancer   34.118.232.129   34.134.125.235   80:32702/TCP   8m53s
PyTorch-gpu-worker-0   ClusterIP   None            <none>         23456/TCP    132m
PyTorch-gpu-worker-0   ClusterIP   None            <none>         23456/TCP    132m
training-operator   ClusterIP   34.118.239.244   <none>         8080/TCP     147m
slb2250@cloudshell:~/amlc-hw3$ inference (amlc-hw3)$

```



MNIST Digit Prediction

6.png



MNIST Digit Prediction

No file chosen
Predicted Digit: 6
[Try another image](#)