

IDS 568 MLOps - Milestone 1: Submission Evidence

Student Name: Kiran Kumar Srinivasa

Course: IDS 568 MLOps

Date: February 1, 2026

1. Live Deployment URLs

- **Cloud Run (FastAPI Container):**
<https://mlops-api-385501609124.us-central1.run.app>
- **Cloud Function (Serverless):**
<https://us-central1-mlops-milestone1-486120.cloudfunctions.net/mlops-function>

2. Testing Evidence & Expected Outputs

A. Cloud Run (Container)

Command:

Bash

```
curl -X POST "https://mlops-api-385501609124.us-central1.run.app/predict" \  
-H "Content-Type: application/json" \  
-d '{"features": [8.3252, 41.0, 6.9841, 1.0238, 322.0, 2.5556, 37.88, -122.23]}'
```

Expected Output:

JSON

```
{  
  "prediction": 4.151938398277785,  
  "model_version": "1.0.0"  
}
```

B. Cloud Function (Serverless)

Command:

Bash

```
curl -X POST "https://us-central1-mlops-milestone1-486120.cloudfunctions.net/mlops-function" \
-H "Content-Type: application/json" \
-d '{"features": [8.3252, 41.0, 6.9841, 1.0238, 322.0, 2.5556, 37.88, -122.23]}'
```

Expected Output:

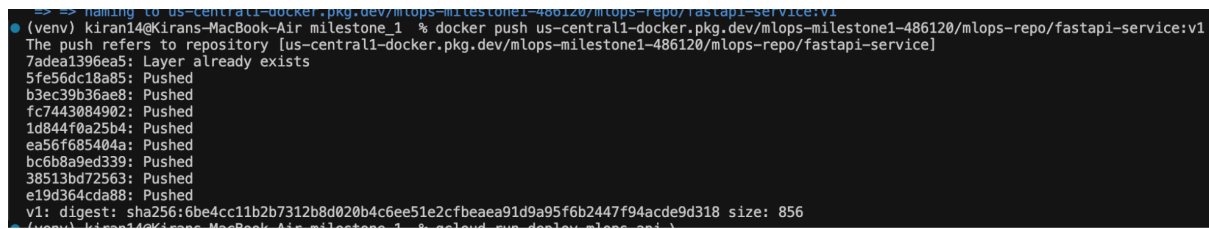
JSON

```
{
  "backend": "Google Cloud Function",
  "model_version": "1.0.0",
  "prediction": 4.151938398277778
}
```

3. Screenshot Evidence

1. Docker Push to Artifact Registry

Evidence of successful container build and push to Google Artifact Registry.

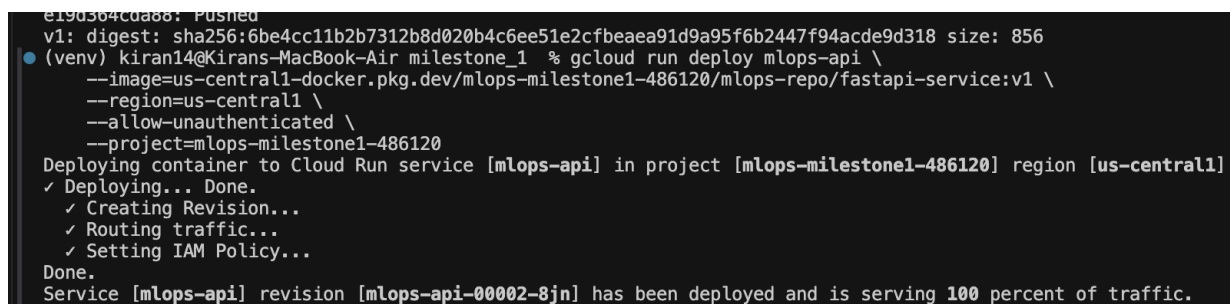


```
➦ ➦ naming to us-central1-docker.pkg.dev/mlops-milestone1-486120/mlops-repo/fastapi-service:v1
(venv) kiran14@Kirans-MacBook-Air milestone_1 % docker push us-central1-docker.pkg.dev/mlops-milestone1-486120/mlops-repo/fastapi-service:v1
The push refers to repository [us-central1-docker.pkg.dev/mlops-milestone1-486120/mlops-repo/fastapi-service]
7adea1396ea5: Layer already exists
5fe56dc18a85: Pushed
b3ec39b36ae8: Pushed
fc7443084902: Pushed
1d844f0a25b4: Pushed
ea56f685404a: Pushed
bc6b8a9ed339: Pushed
38513bd72563: Pushed
e19d364cda88: Pushed
v1: digest: sha256:6be4cc11b2b7312b8d020b4c6ee51e2cfbeaea91d9a95f6b2447f94acde9d318 size: 856
(venv) kiran14@Kirans-MacBook-Air milestone_1 %
```

(This screenshot shows the layers being pushed to `us-central1-docker.pkg.dev`)

2. Cloud Run Deployment Success

Evidence of the service being deployed and serving traffic.



```
e19d364cda88: Pushed
v1: digest: sha256:6be4cc11b2b7312b8d020b4c6ee51e2cfbeaea91d9a95f6b2447f94acde9d318 size: 856
(venv) kiran14@Kirans-MacBook-Air milestone_1 % gcloud run deploy mlops-api \
  --image=us-central1-docker.pkg.dev/mlops-milestone1-486120/mlops-repo/fastapi-service:v1 \
  --region=us-central1 \
  --allow-unauthenticated \
  --project=mlops-milestone1-486120
Deploying container to Cloud Run service [mlops-api] in project [mlops-milestone1-486120] region [us-central1]
✓ Deploying... Done.
✓ Creating Revision...
✓ Routing traffic...
✓ Setting IAM Policy...
Done.
Service [mlops-api] revision [mlops-api-00002-8jn] has been deployed and is serving 100 percent of traffic.
```

(This screenshot shows "Service [mlops-api] ... has been deployed and is serving 100 percent of traffic")

3. Cloud Run Prediction Test

Evidence of a working `curl` request to the Cloud Run endpoint.

```
Setting IAM Policy...
Done.
Service [mlops-api] revision [mlops-api-00002-8jn] has been deployed and is serving 100 percent of traffic.
Service URL: https://mlops-api-385501609124.us-central1.run.app
• (venv) kiran14@Kirans-MacBook-Air milestone_1 % curl -X POST "https://mlops-api-385501609124.us-central1.run.app/predict" \
  -H "Content-Type: application/json" \
  -d '{"features": [8.3252, 41.0, 6.9841, 1.0238, 322.0, 2.5556, 37.88, -122.23]}'
{"prediction":4.151938398277785,"model_version":"1.0.0"}
○ (venv) kiran14@Kirans-MacBook-Air milestone_1 %
```

(This screenshot shows the command and the JSON response with prediction 4.15...)

4. Cloud Function Deployment Details

Evidence of the Cloud Function active state and configuration.

```
--runtime=python311 \ ...
serviceConfig:
  allTrafficOnLatestRevision: true
  availableCpu: '0.1666'
  availableMemory: 256M
  environmentVariables:
    LOG_EXECUTION_ID: 'true'
  ingressSettings: ALLOW_ALL
  maxInstanceCount: 12
  maxInstanceRequestConcurrency: 1
  revision: mlops-function-00002-8jn
  service: projects/mlops-milestone1-486120/locations/us-central1/services/mlops-function
  serviceAccountEmail: 385501609124-compute@developer.gserviceaccount.com
  timeoutSeconds: 60
  uri: https://mlops-function-whzkjna2xq-uc.a.run.app
state: ACTIVE
updateTime: '2026-02-01T21:53:00.480314687Z'
url: https://us-central1-mlops-milestone1-486120.cloudfunctions.net/mlops-function
```

(This screenshot shows the function state: `ACTIVE` and the https trigger URL)

5. Cloud Function Prediction Test

Evidence of a working `curl` request to the Cloud Function endpoint.

```
state: ACTIVE
updateTime: '2026-02-01T21:53:00.480314687Z'
url: https://us-central1-mlops-milestone1-486120.cloudfunctions.net/mlops-function
• (venv) kiran14@Kirans-MacBook-Air function_deploy % curl -m 70 -X POST https://us-central1-mlops-milestone1-486120.cloudfunctions.net/mlops-function \
  -H "Content-Type: application/json" \
  -d '{"features": [8.3252, 41.0, 6.9841, 1.0238, 322.0, 2.5556, 37.88, -122.23]}'
{"backend":"Google Cloud Function","model_version":"1.0.0","prediction":4.151938398277785}
○ (venv) kiran14@Kirans-MacBook-Air function_deploy %
```

(This screenshot shows the command and the response `{"backend": "Google Cloud Function"...`)

Verification

```
📁 Found files in current directory.
=====
🛡️ IDS 568 AUTOMATED SANITY CHECK
=====

[1] Checking Required Files...
✅ main.py found
✅ requirements.txt found
✅ model.pkl found

[2] Checking Dependency Pinning...
✅ All dependencies are pinned

[3] Checking Pydantic Schemas...
✅ Pydantic models found

[4] Testing Cloud Run Endpoint...
✅ Cloud Run success (200 OK)

[5] Testing Cloud Function Endpoint...
✅ Cloud Function success (200 OK)

=====
🎉 Sanity Check Complete!
=====
```