

# MLOPS ASSIGNMENT - 2

## GROUP NO - 45

Name	ID	Contribution
NITESH KUMAR	2024AA05143	100%
LAWLESH KUMAR	2024AA05149	100%
VAIBHAV SAREEN	2024AA05923	100%
VIVEK TRIVEDI	2024AA05922	100%
BINDU MANOJ	2024AA05979	100%

## GITHUB REPO LINKS

**Code Repository :**

[https://github.com/cseniteshkumar/bitsMtech\\_MLOps\\_Assignment\\_2](https://github.com/cseniteshkumar/bitsMtech_MLOps_Assignment_2)

**Github Actions :** [https://github.com/cseniteshkumar/bitsMtech\\_MLOps\\_Assignment\\_2/actions](https://github.com/cseniteshkumar/bitsMtech_MLOps_Assignment_2/actions)

**Docker Link :**

<https://hub.docker.com/repository/docker/cseniteshkumar/mlopsassignment2>

<https://hub.docker.com/r/cseniteshkumar/mlopsassignment2>

**Kubernetes Load Balancer :**

[https://github.com/cseniteshkumar/bitsMtech\\_MLOps\\_Assignment\\_2/blob/main/k8s/service.yaml](https://github.com/cseniteshkumar/bitsMtech_MLOps_Assignment_2/blob/main/k8s/service.yaml)

**Kubernetes / Docker Deployment :**

[https://github.com/cseniteshkumar/bitsMtech\\_MLOps\\_Assignment\\_2/blob/main/k8s/deployment.yaml](https://github.com/cseniteshkumar/bitsMtech_MLOps_Assignment_2/blob/main/k8s/deployment.yaml)

**Model :** [https://github.com/cseniteshkumar/bitsMtech\\_MLOps\\_Assignment\\_2/tree/main/artifacts](https://github.com/cseniteshkumar/bitsMtech_MLOps_Assignment_2/tree/main/artifacts)

**Model Training :**

[https://github.com/cseniteshkumar/bitsMtech\\_MLOps\\_Assignment\\_2/blob/main/modelTraining.ipynb](https://github.com/cseniteshkumar/bitsMtech_MLOps_Assignment_2/blob/main/modelTraining.ipynb)

## YOUTUBE VIDEO LINK

<https://youtu.be/QaEFgxKoqPQ>

# Google Drive LINK

[https://drive.google.com/drive/folders/1\\_I15mlcsOb-yQLkXRd-C0DB1whvUiaA?usp=sharing](https://drive.google.com/drive/folders/1_I15mlcsOb-yQLkXRd-C0DB1whvUiaA?usp=sharing)

## 1. Setup and Installation Instructions

## Local Environment

## Clone the Repository:

Bash

```
git clone <your-repo-link>
```

```
cd bitsMtech MLops Assignment 2
```

## **Environment Setup:**

Bash

```
python3 -m venv .venv
```

```
source .venv/bin/activate # Windows: .venv\Scripts\activate
```

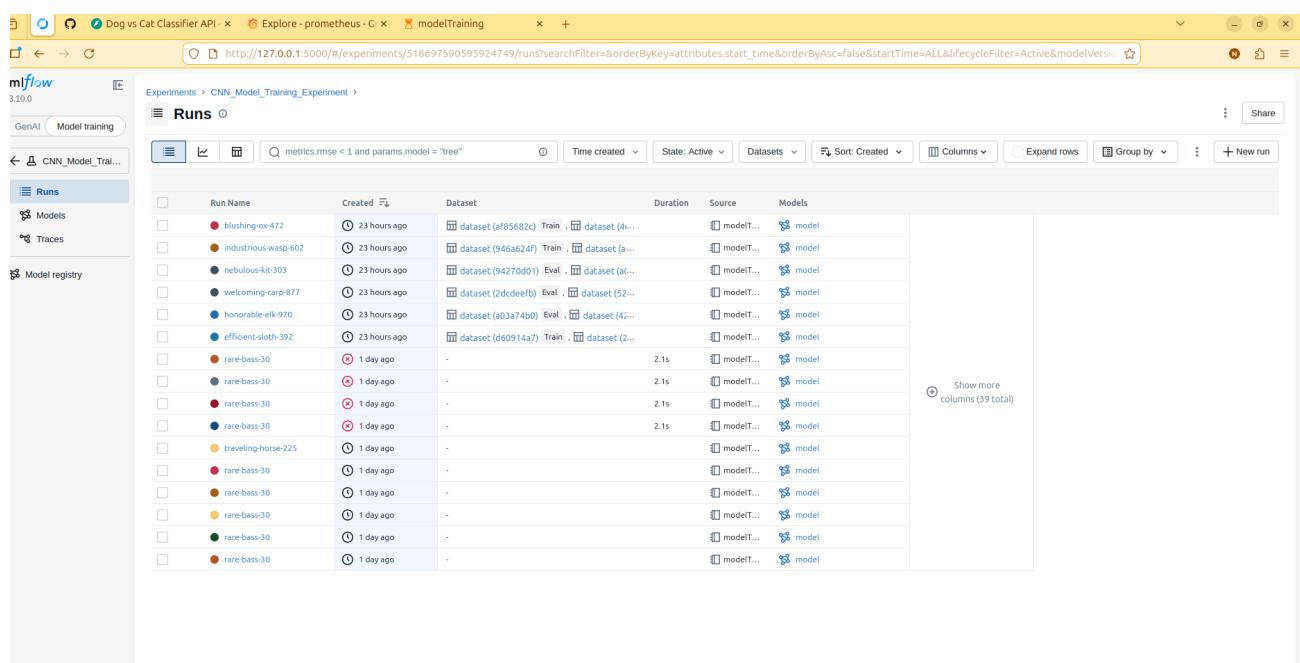
```
pip install -r requirements.txt
```

## Docker Execution

```
docker pull cseniteshkumar/mlopsassignment2:latest
```

```
docker run -p 8000:8000 cseniteshkumar/mllopsassignment2:latest
```

## 2. Experiment Tracking Summary



Screenshot of the miflow interface showing the 'Models' page for the 'CNN\_Model\_Training\_Experiment'. The page displays a table of model attributes, including Model name, Status, Created, Logged from, Source run, Registered models, and Dataset.

Model name	Status	Created	Logged from	Source run	Registered models	Dataset
model	Pending	23 hours ago	modelTraining.ipynb	blushing-ox-472	-	-
model	Pending	23 hours ago	modelTraining.ipynb	industrious-wasp-602	-	-
model	Pending	23 hours ago	modelTraining.ipynb	nebulous-kit-303	-	-
model	Pending	23 hours ago	modelTraining.ipynb	welcoming-carp-877	-	-
model	Pending	23 hours ago	modelTraining.ipynb	honorable-elk-970	-	-
model	Pending	23 hours ago	modelTraining.ipynb	efficient-sloth-392	-	-
model	Pending	1 day ago	modelTraining.ipynb	rare-bass-30	-	-
model	Pending	1 day ago	modelTraining.ipynb	rare-bass-30	-	-
model	Pending	1 day ago	modelTraining.ipynb	rare-bass-30	-	-
model	Pending	1 day ago	modelTraining.ipynb	rare-bass-30	-	-
model	Pending	1 day ago	modelTraining.ipynb	rare-bass-30	-	-
model	Pending	1 day ago	modelTraining.ipynb	rare-bass-30	-	-
model	Pending	1 day ago	modelTraining.ipynb	rare-bass-30	-	-
model	Pending	1 day ago	modelTraining.ipynb	rare-bass-30	-	-
model	Pending	1 day ago	modelTraining.ipynb	rare-bass-30	-	-

Screenshot of the miflow interface showing the 'Runs' page for the 'CNN\_Model\_Training\_Experiment'. The page displays the details of a specific run named 'blushing-ox-472', including Overview, Model metrics, System metrics, Traces, and Artifacts tabs.

**Overview:** No metrics recorded.

**Parameters (24):**

Parameter	Value
opt_clipnorm	None
validation_split	0.0
opt_nesterov	True
opt_weight_decay	None
opt_momentum	0.9
opt_ema_overwrite_frequency	None
opt_gradient_accumulation_steps	None
opt_ema_momentum	0.99
epochs	20
opt_global_clipnorm	None
opt_learning_rate	0.00999999776482582
validation_batch_size	None
shuffle	True
opt_name	SGD

**Logged models (1):** blushing-ox-472

**Source:** modelTraining.ipynb

**Registered prompts:** -

**Datasets:** dataset (af85682c) Train +1

**Tags:** Add tags

**Registered models:** None

### 3. Training SS

Jupyter modelTraining Last Checkpoint: 9 minutes ago

```
# 3. Clean Enumerated Loop
for i, idx in enumerate(random_indices):
    img = Image.open(train_image_list[idx])
    axes[i].imshow(img)
    axes[i].axis('off') # Turn off axis for each specific subplot
    # Optional: axes[i].set_title(f"Index: {idx}", fontsize=8)

plt.tight_layout() # Automatically adjusts spacing
plt.show()
```

[9]:

```
from pathlib import Path
from sklearn.model_selection import train_test_split
```

Jupyter modelTraining Last Checkpoint: 10 minutes ago

```
out = self.relu(self.fc1(out))
out = self.dropout(out) # Apply dropout during forward pass
out = self.fc2(out)
return out

model = Cnn().to(device)
```

[15]:

```
model = Cnn().to(device)
model.train()

[15]:
```

```
Cnn(
    (layer1): Sequential(
        (0): Conv2d(3, 16, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1))
        (1): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (2): ReLU()
        (3): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
    )
    (layer2): Sequential(
        (0): Conv2d(16, 32, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1))
        (1): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (2): ReLU()
        (3): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
    )
    (layer3): Sequential(
        (0): Conv2d(32, 64, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1))
        (1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (2): ReLU()
    )
    (adaptive_pool): AdaptiveAvgPool2d(output_size=(1, 1))
    (fc1): Linear(in_features=64, out_features=10, bias=True)
    (dropout): Dropout(p=0.5, inplace=False)
    (fc2): Linear(in_features=10, out_features=2, bias=True)
    (relu): ReLU()
)
```

[16]:

```
# 1. Modern Optimizer (AdamW is generally preferred over Adam for better weight decay)
optimizer = optim.AdamW(model.parameters(), lr=0.001)

# 2. Loss Function (Standard for classification)
criterion = nn.CrossEntropyLoss()

# 3. Learning Rate Scheduler (Optional but recommended for "latest" performance)
# Reduces the learning rate if the validation loss stops improving
scheduler = optim.lr_scheduler.ReduceLROnPlateau(optimizer, 'min', patience=2)

print(f"Model ready on {device}. Hyperparameters set.")
```

```

# ... EARLY STOPPING ...
if avg_val_loss < best_val_loss - min_delta:
    best_val_loss = avg_val_loss
    early_stop_counter = 0
else:
    early_stop_counter += 1
    if early_stop_counter >= patience:
        print(f"Early stopping at epoch {epoch+1}")
        break
Epoch [1/50]
Train Loss: 0.6666 | Train Acc: 0.5886
Val Loss: 0.6389 | Val Acc: 0.6286

Epoch [2/50]
Train Loss: 0.6355 | Train Acc: 0.6306
Val Loss: 0.6070 | Val Acc: 0.6848

Epoch [3/50]
Train Loss: 0.6103 | Train Acc: 0.6684
Val Loss: 0.6995 | Val Acc: 0.5828

Epoch [4/50]
Train Loss: 0.5945 | Train Acc: 0.6863
Val Loss: 0.5718 | Val Acc: 0.7074

Epoch [5/50]
Train Loss: 0.5850 | Train Acc: 0.6888
Val Loss: 0.5866 | Val Acc: 0.6838

Epoch [6/50]
Train Loss: 0.5802 | Train Acc: 0.6963
Val Loss: 0.5507 | Val Acc: 0.7190

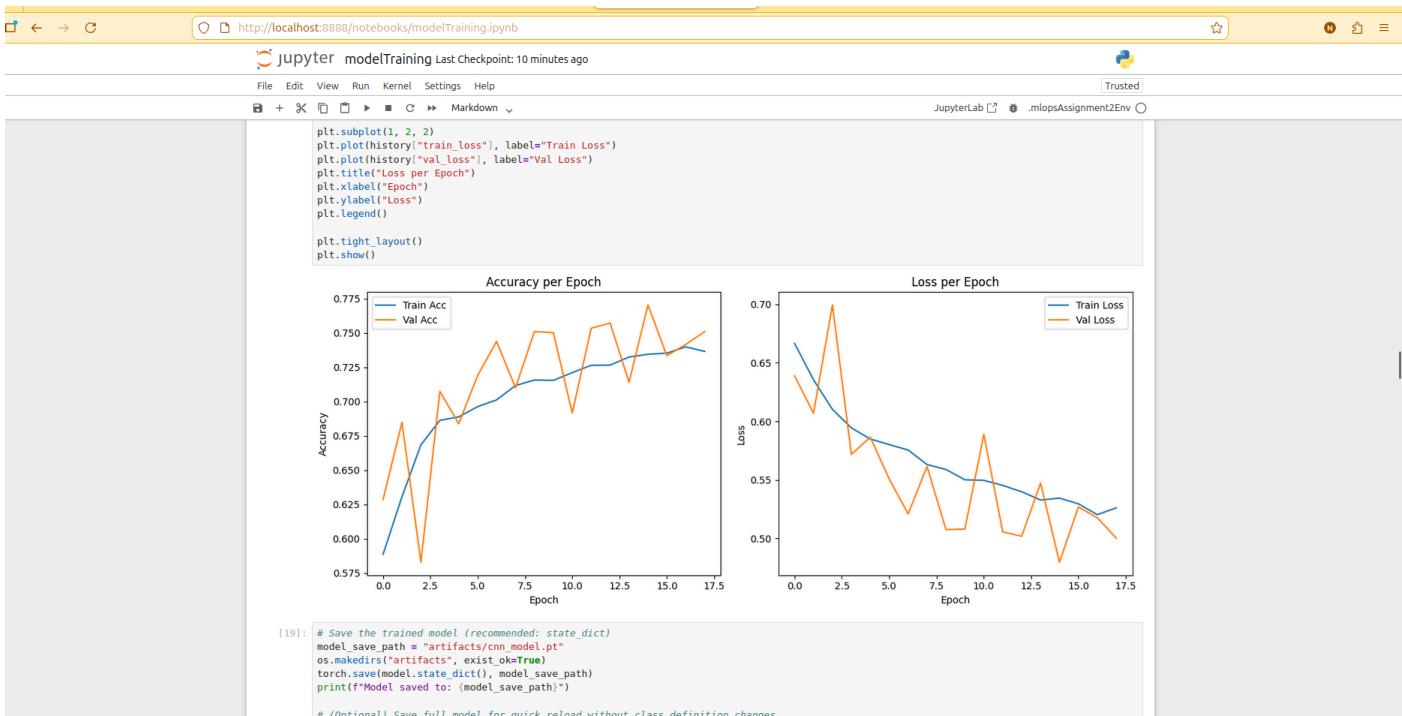
Epoch [7/50]
Train Loss: 0.5756 | Train Acc: 0.7012
Val Loss: 0.5209 | Val Acc: 0.7438

Epoch [8/50]
Train Loss: 0.5631 | Train Acc: 0.7117
Val Loss: 0.5614 | Val Acc: 0.7100

Epoch [9/50]
Train Loss: 0.5589 | Train Acc: 0.7157
Val Loss: 0.5076 | Val Acc: 0.7510

Epoch [10/50]

```



## 4. CI/CD and Deployment Workflow

### CI (Continuous Integration)



The screenshot shows the GitHub Actions CI Pipeline configuration in a VS Code editor. The left sidebar displays the project structure, including files like `requirements_pinned.txt`, `deployment.yaml`, `.gitignore`, and `cyaml`. The main editor area shows the `cyaml` file content:

```
name: CI Pipeline
on:
  push:
    branches: [ main, develop ]
  pull_request:
    branches: [ main, develop ]
jobs:
  test:
    runs-on: ubuntu-latest
    strategy:
      matrix:
        python-version: ["3.9", "3.10"]
    steps:
      - uses: actions/checkout@v3
      - uses: actions/setup-python@v4
        with:
          python-version: ${{ matrix.python-version }}
          cache: 'pip'
      - run:
          pip install --upgrade pip
          pip install pytest pytest-cov -r requirements_fastapi.txt
      - run:
          pytest tests/ -v --cov=app --cov-report=xml
      - uses: codecov/codecov-action@v3
        with:
          file: ./coverage.xml
          fail_ci_if_error: false
```

The screenshot shows the GitHub Actions CI Pipeline page for the repository 'bitsMtech\_MLOps\_Assignment\_2'. The sidebar on the left lists various CI Pipeline management options like Management, Caches, Attestations, Runners, Usage metrics, and Performance metrics. The main content area displays a table of 24 workflow runs, each with a status indicator (green checkmark), name, commit information, branch (main), and timestamp. The runs are ordered by creation time, with the most recent at the top.

Workflow Run	Commit	Branch	Status	Timestamp
Train_Log_Generation	Commit 944d5fd1 pushed by cseniteshkumar	main	In progress	now
update Git Ignore	Commit 947e76d pushed by cseniteshkumar	main	Success	14 minutes ago
update gitIgnore	Commit 035a581 pushed by cseniteshkumar	main	Success	14 minutes ago
k8s	Commit b911ed1 pushed by cseniteshkumar	main	Success	26 minutes ago
updateReadMeFile	Commit a207b73 pushed by cseniteshkumar	main	Success	42 minutes ago
pro update	Commit 08ba414 pushed by cseniteshkumar	main	Success	Today at 9:46 PM
promethesia and graphana update	Commit 271f818 pushed by cseniteshkumar	main	Success	Today at 9:30 PM
promethesia update	Commit 740162e pushed by cseniteshkumar	main	Success	Today at 9:18 PM

github.com/cseniteshkumar/bitsMtech\_MLOps\_Assignment\_2/actions/runs/22281441735

All Bookmarks

cseñiteshkumar / bitsMtech\_MLOps\_Assignment\_2

Code Issues Pull requests Agents Actions Projects Wiki Security Insights Settings

CI Pipeline

update Git Ignore #23

Summary

All jobs

Triggered via push 14 minutes ago

cseñiteshkumar pushed → 947e7ed main

Status Success Total duration 2m 29s Artifacts

cl.yml on: push

Matrix: test

test (3.10) 2m 25s

test (3.9) 2m 2s

Run details Usage Workflow file

Re-run all jobs

github.com/cseniteshkumar/bitsMtech\_MLOps\_Assignment\_2/actions/runs/22281441735/job/64452611656

All Bookmarks

cseñiteshkumar / bitsMtech\_MLOps\_Assignment\_2

Code Issues Pull requests Agents Actions Projects Wiki Security Insights Settings

CI Pipeline

update Git Ignore #23

Summary

All jobs

test (3.9)

test (3.10)

Run details Usage Workflow file

Re-run all jobs

test (3.10)

succeeded 12 minutes ago in 2m 25s

Search logs

- > Set up job 2s
- > Run actions/checkout@v3 2s
- > Run actions/setup-python@v4 1s
- > Run pip install --upgrade pip 1m 50s
- > Run pytest tests/-v --cov=app --cov-report=xml 7s
- > Run codecov/codecov-action@v3 2s
- > Post Run actions/setup-python@v4 18s
- > Post Run actions/checkout@v3 1s
- > Complete job 0s

1 Cleaning up orphan processes

1 Cleaning up orphan processes

CI Pipeline

update Git Ignore #23

Re-run all jobs

Summary

All jobs

test (3.9)

succeeded 13 minutes ago in 2m 2s

Search logs

test (3.9)

test (3.10)

Run details

Usage

Workflow file

test (3.9)

Set up job 1s

Run actions/checkout@v3 2s

Run actions/setup-python@v4 3s

Run pip install --upgrade pip 1m 14s

Run pytest tests/-v --cov=app --cov-report=xml 5s

Run codecov/codecov-action@v3 2s

Post Run actions/setup-python@v4 0s

Post Run actions/checkout@v3 0s

Complete job 0s

Summary

All jobs

test (3.9)

succeeded 13 minutes ago in 2m 2s

Search logs

test (3.9)

test (3.10)

Run details

Usage

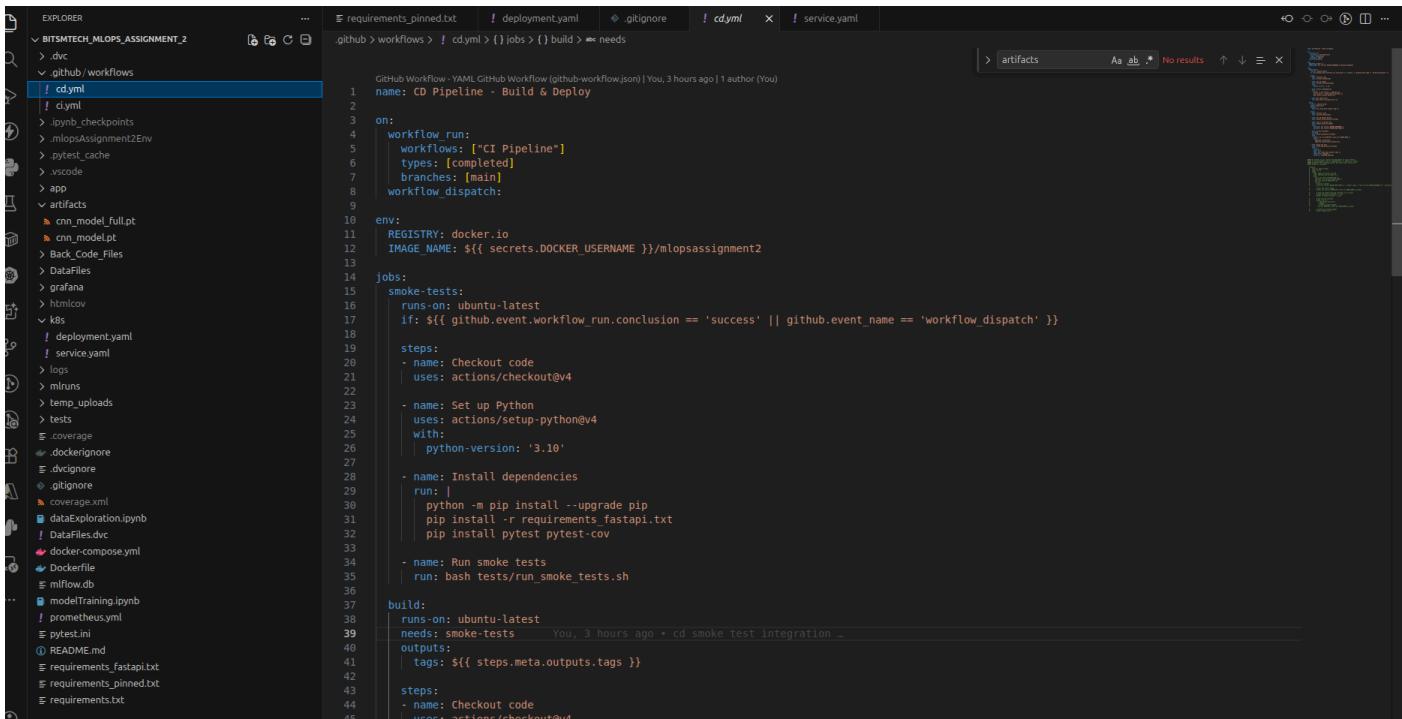
Workflow file

test (3.9)

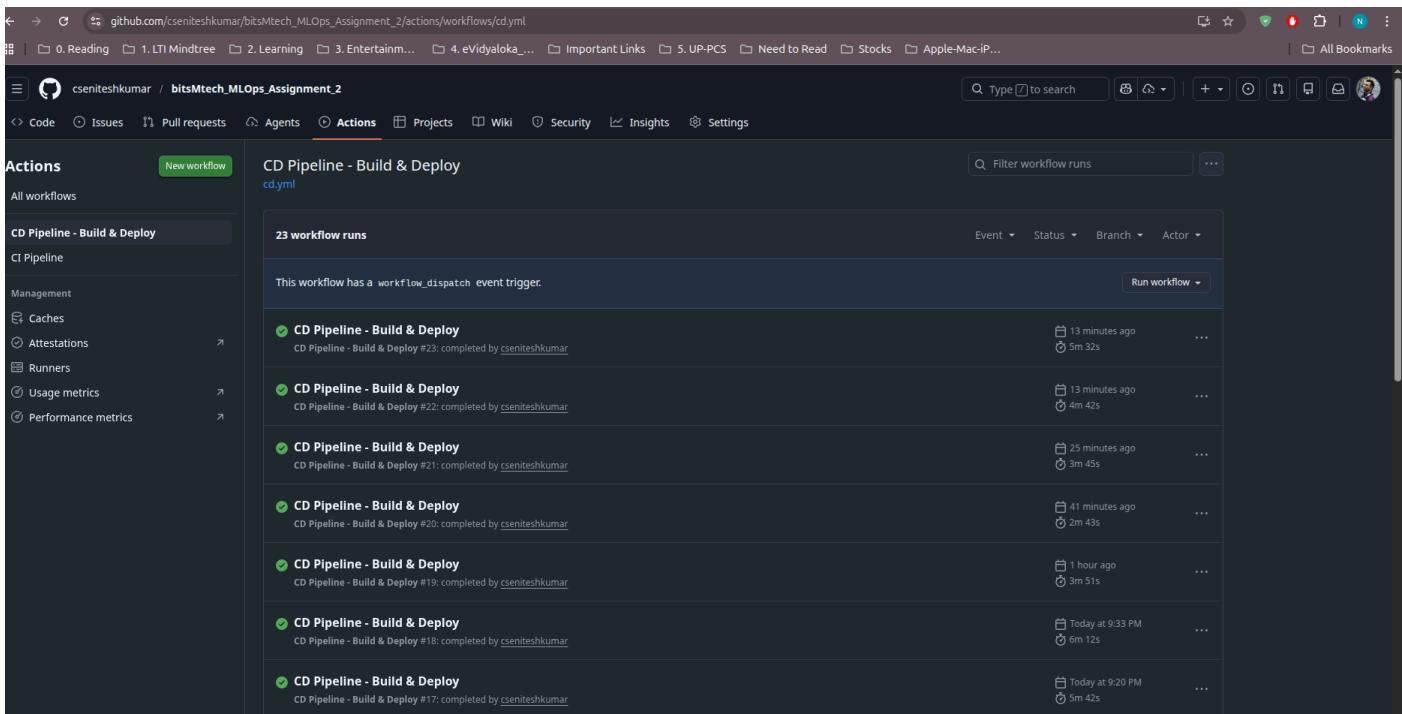
Run pytest tests/-v --cov=app --cov-report=xml

```
13 cacheoir: pytest_cache
14 rootdir: /home/runner/work/bitsMtech_MLOps_Assignment_2/bitsMtech_MLOps_Assignment_2
15 configfile: pytest.ini
16 plugins: cov-7.0.0, anyio-4.12.1
17 collecting ... collected 18 items
18
19 tests/test_deployment.py::TestDeployment::test_app_startup PASSED [ 5%]
20 tests/test_deployment.py::TestDeployment::test_health_check PASSED [ 11%]
21 tests/test_deployment.py::TestDeployment::test_api_routes_exist PASSED [ 16%]
22 tests/test_deployment.py::TestDeployment::test_prediction_endpoint_exists PASSED [ 22%]
23 tests/test_deployment.py::TestDeployment::test_valid_prediction_request PASSED [ 27%]
24 tests/test_deployment.py::TestDeployment::test_invalid_prediction_request PASSED [ 33%]
25 tests/test_deployment.py::TestDeployment::test_empty_request PASSED [ 38%]
26 tests/test_deployment.py::TestDeployment::test_model_inference_consistency PASSED [ 44%]
27 tests/test_deployment.py::TestDeployment::test_response_format PASSED [ 50%]
28 tests/test_deployment.py::TestDeployment::test_concurrent_requests PASSED [ 55%]
29 tests/test_deployment.py::TestDeployment::test_large_input PASSED [ 61%]
30 tests/test_deployment.py::TestDeployment::test_special_values PASSED [ 66%]
31 tests/test_deployment.py::TestDeployment::test_model_loaded PASSED [ 72%]
32 tests/test_deployment.py::TestDeployment::test_config_loaded PASSED [ 77%]
33 tests/test_deployment.py::TestDeployment::test_response_headers PASSED [ 83%]
34 tests/test_deployment.py::TestErrorHandling::test_malformed_json PASSED [ 88%]
35 tests/test_deployment.py::TestErrorHandling::test_wrong_content_type PASSED [ 94%]
36 tests/test_deployment.py::TestErrorHandling::test_missing_required_fields PASSED [100%]
37
38 ===== warnings summary =====
39 tests/test_deployment.py::TestDeployment::test_malformed_json
40 tests/test_deployment.py::TestErrorHandling::test_wrong_content_type
41 /opt/hostedtoolcache/Python/3.9.25/x64/lib/python3.9/site-packages/httpx/_models.py:408: DeprecationWarning: Use 'content=<...>' to upload raw bytes/text content.
42     headers, stream = encode_request()
43
44 -- Docs: https://docs.pytest.org/en/stable/how-to/capture-warnings.html
45 ===== tests coverage =====
46 coverage: platform linux, python 3.9.25-final-0
47
48 Coverage XML written to file coverage.xml
49 ===== 18 passed, 2 warnings in 0.18s =====
```

## CD (Continuous Deployment)



```
version: '3.8'
name: CD Pipeline - Build & Deploy
on:
  workflow_dispatch:
  workflow_run:
    workflows: ['CI Pipeline']
    types: [completed]
    branches: [main]
  jobs:
    smoke-tests:
      runs-on: ubuntu-latest
      if: ${{ github.event.workflow.run.conclusion == 'success' || github.event_name == 'workflow_dispatch' }}
      steps:
        - name: Checkout code
          uses: actions/checkout@v4
        - name: Set up Python
          uses: actions/setup-python@v4
          with:
            python-version: '3.10'
        - name: Install dependencies
          run:
            python -m pip install --upgrade pip
            pip install -r requirements_fastapi.txt
            pip install pytest pytest-cov
        - name: Run smoke tests
          run: bash tests/run_smoke_tests.sh
      build:
        runs-on: ubuntu-latest
        needs: smoke-tests
        outputs:
          tags: ${{ steps.meta.outputs.tags }}
      steps:
        - name: Checkout code
          uses: actions/checkout@v4
```



The screenshot shows the GitHub Actions interface for the repository 'bitsMtech.MLOps\_Assignment\_2'. The 'Actions' tab is selected. On the left, there's a sidebar with navigation links like 'Code', 'Issues', 'Pull requests', 'Agents', 'Actions', 'Projects', 'Wiki', 'Security', 'Insights', and 'Settings'. The 'Actions' section has a 'New workflow' button and a 'CD Pipeline - Build & Deploy' card. The card displays the workflow name, a 'cd.yml' file link, and a 'Filter workflow runs' search bar. Below the card is a table titled '23 workflow runs' with columns for 'Event', 'Status', 'Branch', and 'Actor'. The table lists 23 completed runs, each with a green circular icon, the run ID, the actor (cseniteshkumar), and the timestamp. For example, the most recent run was completed 13 minutes ago.

Event	Status	Branch	Actor
13 minutes ago	Completed	main	cseniteshkumar
13 minutes ago	Completed	main	cseniteshkumar
25 minutes ago	Completed	main	cseniteshkumar
41 minutes ago	Completed	main	cseniteshkumar
1 hour ago	Completed	main	cseniteshkumar
Today at 9:33 PM	Completed	main	cseniteshkumar
Today at 9:20 PM	Completed	main	cseniteshkumar
13 minutes ago	Completed	main	cseniteshkumar
13 minutes ago	Completed	main	cseniteshkumar
25 minutes ago	Completed	main	cseniteshkumar
41 minutes ago	Completed	main	cseniteshkumar
1 hour ago	Completed	main	cseniteshkumar
Today at 9:33 PM	Completed	main	cseniteshkumar
Today at 9:20 PM	Completed	main	cseniteshkumar
13 minutes ago	Completed	main	cseniteshkumar
13 minutes ago	Completed	main	cseniteshkumar
25 minutes ago	Completed	main	cseniteshkumar
41 minutes ago	Completed	main	cseniteshkumar
1 hour ago	Completed	main	cseniteshkumar
Today at 9:33 PM	Completed	main	cseniteshkumar
Today at 9:20 PM	Completed	main	cseniteshkumar
13 minutes ago	Completed	main	cseniteshkumar
13 minutes ago	Completed	main	cseniteshkumar
25 minutes ago	Completed	main	cseniteshkumar
41 minutes ago	Completed	main	cseniteshkumar
1 hour ago	Completed	main	cseniteshkumar
Today at 9:33 PM	Completed	main	cseniteshkumar
Today at 9:20 PM	Completed	main	cseniteshkumar

github.com/cseniteshkumar/bitsMtech\_MLOps\_Assignment\_2/actions/runs/22281482763

All Bookmarks

cseñiteshkumar / bitsMtech\_MLOps\_Assignment\_2

Code Issues Pull requests Agents Actions Projects Wiki Security Insights Settings

CD Pipeline - Build & Deploy

CD Pipeline - Build & Deploy #23

Re-run all jobs ...

Summary

All jobs

smoke-tests build

Run details Usage Workflow file

Triggered via workflow run 13 minutes ago

cseñiteshkumar completed → 947e76d Status Success Total duration 5m 32s Artifacts

cd.yml on: workflow\_run

smoke-tests 2m 5s build 3m 20s

Annotations

1 warning

build Skip output 'tags' since it may contain secret.

This screenshot shows the GitHub Actions pipeline summary for a specific run. It displays the status as 'Success' with a total duration of 5m 32s. The pipeline consists of two jobs: 'smoke-tests' (2m 5s) and 'build' (3m 20s). There is one annotation warning about skipping output 'tags' due to potential secrets. The sidebar on the left shows other repository details like '0. Reading' and '1. LTI Mindtree'.

github.com/cseniteshkumar/bitsMtech\_MLOps\_Assignment\_2/actions/runs/22281482763/job/64452715499

All Bookmarks

cseñiteshkumar / bitsMtech\_MLOps\_Assignment\_2

Code Issues Pull requests Agents Actions Projects Wiki Security Insights Settings

CD Pipeline - Build & Deploy

CD Pipeline - Build & Deploy #23

Re-run all jobs ...

Summary

All jobs

smoke-tests build

Run details Usage Workflow file

smoke-tests succeeded 11 minutes ago in 2m 5s

Set up job 0s

Checkout code 3s

Set up Python 0s

Install dependencies 1m 49s

Run smoke tests 10s

Post Set up Python 0s

Post Checkout code 1s

Complete job 0s

Search logs

This screenshot shows the detailed log for the 'smoke-tests' job from the previous run. The job succeeded in 2m 5s. It includes steps for setting up the job, checking out code, setting up Python, installing dependencies, running smoke tests, and completing the job. The log also features a search bar for logs.

github.com/cseniteshkumar/bitsMtech\_MLOps\_Assignment\_2/actions/runs/22281482763/job/64452715499

All jobs

**smoke-tests**

succeeded 11 minutes ago in 2m 5s

**Run Smoke tests**

```

40 tests/test_smoke.py::test_smoke_tests::test_1b_cuda_device_if_available SKIPPED [ 2%]
41 tests/test_smoke.py::TestSmokeTests::test_16_normalization_parameters PASSED [ 80%]
42 tests/test_smoke.py::TestSmokeTests::test_17_image_transform_output PASSED [ 85%]
43 tests/test_smoke.py::TestSmokeTests::test_18_inference_speed PASSED [ 90%]
44 tests/test_smoke.py::TestSmokeTests::test_19_pytorch_available PASSED [ 95%]
45 tests/test_smoke.py::TestSmokeTests::test_20_pillow_available PASSED [100%]
46
47 ===== 19 passed, 1 skipped in 0.15s =====
48
49 Step 2: Running with coverage...
50 ===== test session starts =====
51 platform linux -- Python 3.10.19, pytest-9.0.2, pluggy-1.6.0
52 rootdir: /home/runner/work/bitsMtech_MLOps_Assignment_2/bitsMtech_MLOps_Assignment_2
53 configfile: pytest.ini
54 plugins: cov-7.6.0, anyio-4.12.1
55 collected 20 items
56
57 tests/test_smoke.py ..... [100%]
58
59 ===== tests coverage =====
60 coverage: platform linux, python 3.10.19-final-0
61
62      Name           Stmts   Miss  Cover   Missing
63 -----
64 app/__init__.py          0     0  100%
65 app/api/routes.py       34    21  38%  13-20, 37-68, 76
66 app/cors/config.py      16     3  81%  29-22
67 app/main.py              15     3  80%  36, 43-44
68 app/models/model.py     58     0  100%
69 app/schemas/_init_.py    2     0  100%
70 app/schemas/prediction.py 16     0  100%
71 app/services/inference.py 43    26  40%  17-22, 35-51, 56-60, 73-87
72
73 TOTAL                   184    53  71%
74 ===== 19 passed, 1 skipped in 0.24s =====
75
76 =====
77 Smoke Tests Complete

```

github.com/cseniteshkumar/bitsMtech\_MLOps\_Assignment\_2/actions/runs/22281482763/job/64452809008

All jobs

**build**

succeeded 8 minutes ago in 3m 20s

**Annotations**

1 warning

**build**

Set up job  
Checkout code  
Set up Docker Buildx  
Log in to Docker Hub  
Extract Metadata  
Build and push  
Post Build and push  
Post Log in to Docker Hub  
Post Set up Docker Buildx  
Post Checkout code  
Complete job

github.com/cseniteshkumar/bitsMtech\_MLOps\_Assignment\_2/actions/runs/22281482763/job/64452809008

All jobs

smoke-tests

**build**

Run details

Usage

Workflow file

**Annotations**

1 warning

**build**

succeeded 8 minutes ago in 3m 20s

Search logs

Set up job 5s

Checkout code 3s

Set up Docker Buildx 5s

**Log in to Docker Hub** 1s

- Run docker/login-action@v3
- Logging into docker.io...
- Login Succeeded!

Extract Metadata 1s

Build and push 2m 59s

**Post Build and push** 8s

- Post job cleanup.
- Removing temp folder /home/runner/work/\_temp/docker-actions-toolkit-R0UNba
- Post cache

**Post Log in to Docker Hub** 8s

- Post job cleanup.
- Logout from docker.io
- Post cache

**Post Set up Docker Buildx** 3s

- Post job cleanup.
- Removing builder
- Cleaning up certificates
- Post cache

**Post Checkout code** 1s

```
1 ► Run docker/login-action@v3
9 Logging into docker.io...
10 Login Succeeded!
1 ► Extract Metadata
2 ► Build and push
3 ► Post Build and push
1 Post job cleanup.
2 ► Removing temp folder /home/runner/work/_temp/docker-actions-toolkit-R0UNba
3 ► Post cache
5 ► Post cache
1 Post job cleanup.
2 ► Logout from docker.io
5 ► Post cache
1 Post job cleanup.
2 ► Removing builder
5 ► Cleaning up certificates
6 ► Post cache
1 Post Checkout code
```

github.com/cseniteshkumar/bitsMtech\_MLOps\_Assignment\_2/actions/runs/22281482763/job/64452809008

All jobs

smoke-tests

**build**

Run details

Usage

Workflow file

**CD Pipeline - Build & Deploy**

**CD Pipeline - Build & Deploy #23**

Re-run all jobs

**Annotations**

1 warning

**build**

succeeded 8 minutes ago in 3m 20s

Search logs

Set up job 5s

Checkout code 3s

Set up Docker Buildx 5s

**Log in to Docker Hub** 1s

Extract Metadata 1s

Build and push 2m 59s

**Post Build and push** 8s

- Post job cleanup.
- Removing temp folder /home/runner/work/\_temp/docker-actions-toolkit-R0UNba
- Post cache

**Post Log in to Docker Hub** 8s

- Post job cleanup.
- Logout from docker.io
- Post cache

**Post Set up Docker Buildx** 3s

- Post job cleanup.
- Removing builder
- Cleaning up certificates
- Post cache

**Post Checkout code** 1s

```
1 ► Run docker/login-action@v3
9 Logging into docker.io...
10 Login Succeeded!
1 ► Extract Metadata
2 ► Build and push
3 ► Post Build and push
1 Post job cleanup.
2 ► Removing temp folder /home/runner/work/_temp/docker-actions-toolkit-R0UNba
3 ► Post cache
5 ► Post cache
1 Post job cleanup.
2 ► Logout from docker.io
5 ► Post cache
1 Post job cleanup.
2 ► Removing builder
5 ► Cleaning up certificates
6 ► Post cache
1 Post cache
1 Post Checkout code
```

The screenshot shows the GitHub Actions dashboard for the repository 'bitsMtech\_MLOps\_Assignment\_2'. The left sidebar contains navigation links for Code, Issues, Pull requests, Agents, Actions, Projects, Wiki, Security, Insights, and Settings. The main area displays 'All workflows' and '51 workflow runs'. The runs are listed in descending order of completion time, with the most recent at the top. Each run is represented by a card with a status icon, name, event, status, branch, and duration. A search bar at the top right allows filtering of workflow runs.

## 5. Smoke Test Log

```

niteshkumar@cseniteshkum... ✘ niteshkumar@cseniteshkum... ✘ niteshkumar@cseniteshkum... ✘ niteshkumar@cseniteshkum... ✘ niteshkumar@cseniteshkum... ✘ niteshkumar@cseniteshkum... ✘ niteshkumar@cseniteshkum... ✘
$ cd /media/niteshkumar/SSD_Store_0_nvme/allPythoncodesWithPipEnv/BitsLearning/MLOps_Assignment_2/bitsMtech_MLOps_Assignment_2
$ pytest --cov=bitsMtech_MLOps_Assignment_2 --cov-report=html --cov-branch --log-level=INFO
$ coverage report -m
  Name          Stmts   Miss  Cover   Missing
+----+-----+-----+-----+
  tests/test_deployment.py::TestDeployment::test_app_startup      1       0    100%
  tests/test_deployment.py::TestDeployment::test_health_check     1       0    100%
  tests/test_deployment.py::TestDeployment::test_api_routes_exist 1       0    100%
  tests/test_deployment.py::TestDeployment::test_prediction_endpoint_exists 1       0    100%
  tests/test_deployment.py::TestDeployment::test_valid_prediction_request 1       0    100%
  tests/test_deployment.py::TestDeployment::test_invalid_prediction_request 1       0    100%
  tests/test_deployment.py::TestDeployment::test_empty_request     1       0    100%
  tests/test_deployment.py::TestDeployment::test_model_inference_consistency 1       0    100%
  tests/test_deployment.py::TestDeployment::test_response_format    1       0    100%
  tests/test_deployment.py::TestDeployment::test_concurrent_requests 1       0    100%
  tests/test_deployment.py::TestDeployment::test_large_input        1       0    100%
  tests/test_deployment.py::TestDeployment::test_special_values     1       0    100%
  tests/test_deployment.py::TestDeployment::test_model_loaded       1       0    100%
  tests/test_deployment.py::TestDeployment::test_config_loaded     1       0    100%
  tests/test_deployment.py::TestDeployment::test_response_headers   1       0    100%
  tests/test_deployment.py::TestErrorHandling::test_malformed_json 1       0    100%
  tests/test_deployment.py::TestErrorHandling::test_wrong_content_type 1       0    100%
  tests/test_deployment.py::TestErrorHandling::test_missing_required_fields 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_01_cnn_model_instantiation 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_02_cnn_model_structure 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_03_cnn_forward_pass     1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_04_cnn_batch_inference 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_05_model_checkpoint_exists 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_06_model_checkpoint_loads 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_07_image_classifier_init 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_08_image_classifier_transforms 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_09_image_classifier_invalid_model_path 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_10_inference_with_valid_image 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_11_inference_output_structure 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_12_inference_probability_validity 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_13_inference_with_invalid_image 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_14_cpu_device_placement 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_15_cuda_device_if_available 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_16_normalization_parameters 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_17_image_transform_output 1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_18_inference_speed     1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_19_pytorch_available    1       0    100%
  tests/test_smoke.py::TestSmokeTests::test_20_pillow_available     1       0    100%
Selected 38 items

```

## 6. Fast API with Docker/Kubernetes Deployment

Runs - Experiment 51669 × Dog vs Cat Classifier API - × Explore - prometheus - × +

http://127.0.0.1:8000/docs#

## Dog vs Cat Classifier API 1.0.0 OAS 3.1

/openapi.json

Cat vs Dog Image Classification API

**default**

GET /metrics Metrics

GET / Read Root

**predictions**

GET /api/health Health Check

POST /api/predict Predict Image

GET /api/ Root

**Schemas**

Body\_predict\_image\_api\_predict\_post > Expand all object

HTTPValidationError > Expand all object

HealthResponse > Expand all object

Runs - Experiment 51669 × Dog vs Cat Classifier API - × Explore - prometheus - × +

http://127.0.0.1:8000/docs#/predictions/health\_check\_api\_health\_get

Execute Clear

**Responses**

Curl

```
curl -X 'GET' \
  'http://127.0.0.1:8000/api/health' \
  -H 'accept: application/json'
```

Request URL

http://127.0.0.1:8000/api/health

Server response

Code Details

Undocumented TypeError: NetworkError when attempting to fetch resource.

**Responses**

Code	Description	Links
200	Successful Response	No links

Media type

application/json

Controls Accept header.

Example Value | Schema

```
{ "status": "string", "model_loaded": true }
```

POST /api/predict Predict Image

GET /api/ Root

Runs - Experiment 51669 | Dog vs Cat Classifier API - Explore - prometheus - + http://127.0.0.1:8000/docs#/predictions/predict\_image\_api\_predict\_post

**POST /api/predict Predict Image**

Predict if uploaded image is a cat or dog

Args: file (jpg, jpeg, png)

Returns: Prediction result with confidence scores

**Parameters**

No parameters

**Request body required**

file required  
string(\$binary)

**Responses**

Code	Description	Links
200	Successful Response	No links

Media type application/json  
Controls Accept header.

Example Value | Schema

```
{ "cat_probability": 0.0124, "confidence": 0.9876, "dog_probability": 0.9876, "prediction": "dog", "success": true }
```

Try it out multipart/form-data

Runs - Experiment 51669 | Dog vs Cat Classifier API - Explore - prometheus - + http://127.0.0.1:8000/docs#/predictions/predict\_image\_api\_predict\_post

Returns: Prediction result with confidence scores

**Parameters**

No parameters

**Request body required**

file required  
string(\$binary) No file selected.

**Responses**

Code	Description	Links
200	Successful Response	No links
422	Validation Error	No links

Media type application/json  
Controls Accept header.

Example Value | Schema

```
{ "cat_probability": 0.0124, "confidence": 0.9876, "dog_probability": 0.9876, "prediction": "dog", "success": true }
```

Cancel Reset multipart/form-data

A file upload dialog is displayed, showing a grid of nine image thumbnails. The third thumbnail from the top-left is highlighted with a gray border. The dialog has a title bar "File Upload - Dog vs Cat Classifier API - Swagger UI — P..." and buttons for "Cancel" and "Reset".

The screenshot shows the "Explore" feature of an API documentation tool. A POST request is being made to `http://127.0.0.1:8000/api/predict`. The "Responses" section includes a "Curl" command and a "Request URL". Below this, the "Server response" section shows two entries: a successful response (200) and a validation error (422). The successful response example is a JSON object:

```
{
  "cat_probability": 0.0124,
  "dog_probability": 0.9876,
  "prediction": "dog",
  "success": true
}
```

The validation error example is a JSON object:

```
{
  "detail": ...
}
```

## 7. Docker and Kubernetes

The screenshot shows the Docker Hub interface for the repository `cseniteshkumar/mllopsassignment2`. The repository has 183 stars and was last pushed 1 minute ago. It contains 9 tags. The "General" tab is selected, showing the following table of tags:

Tag	OS	Type	Pulled	Pushed
<code>sha-44d5fd1</code>		Image	less than 1 day	1 minute
<code>latest</code>		Image	less than 1 day	1 minute
<code>sha-947e76d</code>		Image	less than 1 day	16 minutes
<code>sha-b911ed1</code>		Image	less than 1 day	29 minutes
<code>sha-a207b73</code>		Image	less than 1 day	about 1 hour

A sidebar on the left shows the user's profile and various account management options like Billing, Usage, Pulls, and Storage. A right sidebar for "buildcloud" offers Docker Build Cloud services.

hub.docker.com/repository/docker/cseniteshkumar/mlopsassignment2/tags

Hardened Images Add a description

Collaborations Add a category

Settings

Default privacy

Notifications

Billing Usage Pulls Storage

General Tags Image Management Collaborators Webhooks Settings

Sort by Newest Filter tags Delete

TAG sha-44d5fd1 Last pushed 6 minutes by cseniteshkumar

Digest OS/ARCH Last pull Compressed size

457a61a2697a linux/amd64 less than 1 day 3.83 GB

TAG latest Last pushed 6 minutes by cseniteshkumar

Digest OS/ARCH Last pull Compressed size

457a61a2697a linux/amd64 less than 1 day 3.83 GB

TAG sha-947e76d Last pushed 21 minutes by cseniteshkumar

Digest OS/ARCH Last pull Compressed size

b1b674188860 linux/amd64 less than 1 day 3.83 GB

hub.docker.com/u/cseniteshkumar

Explore My Hub Search Docker Hub CtrlK

cseniteshkumar Edit profile  Community User

Repositories Starred

Search by repository name Displaying 1 to 2 of 2 repositories

IMAGE cseniteshkumar/mlopsassignment2 cseniteshkumar

Pulls 193 Stars 0 Last Updated 7 minutes

IMAGE cseniteshkumar/fastapi\_mlopsassignment2 cseniteshkumar

Pulls 17 Stars 0 Last Updated about 4 hours

## While internal system testing

```

niteshkumar@cseniteshkum... x niteshkumar@cseniteshkum... x niteshkumar@cseniteshkum... x niteshkumar@cseniteshkum... x niteshkumar@cseniteshkum... x niteshkumar@cseniteshkum... x niteshkumar@cseniteshkum... x
from app.api.routes import router
[!] Building 1.9s (11/11) FINISHED
> [internal] load Build definition from Dockerfile
>=> transferring dockerfile: 210B
> [internal] load metadata for docker.io/library/python:3.9-slim
> [auth] library/python:pull token for registry-1.docker.io
> [internal] load .dockerignore
>=> transferring context: 1.29kB
> [1/5] FROM docker.io/library/python:3.9-slim@sha256:2d97f6910b16bd338d3060f261f53f144965f755599aab1acd1e13cf1731b1b
>=> resolve docker.io/library/python:3.9-slim@sha256:d97f6910b16bd338d3060f261f53f144965f755599aab1acd1e13cf1731b1b
> [internal] load build context
>=> transferring context: 1.78kB
> CACHED [2/5] WORKDIR /app
> CACHED [3/5] COPY requirements_fastapi.txt .
> CACHED [4/5] RUN pip install --no-cache-dir -r requirements_Fastapi.txt
> [5/5] COPY . .
> exporting to image
>=> exporting layers
>=> exporting manifest sha256:8e494acf293ab288bf2eb046d37dd78ebcd6fbeda45b211e1b319ad1b6ae80b2
>=> exporting config sha256:f04df97af18077d087dff9ec4cabe247a8787712170ce0530179e73fdaf18
>=> exporting attestation manifest sha256:ted2cd44a690b267207fb6901fd96fe95c2f163324d32f6abd979c72067611
>=> exporting manifest list sha256:f25ed79d2bbf9634894e682acd35386789502bb7573467616110664f6af16
>=> naming to docker.io/library/fastapi_mlopsassignment2:latest
>=> unpacking to docker.io/library/fastapi_mlopsassignment2:latest
mlopsAssignment2Env) niteshkumar@cseniteshkum...:~/bitsMtech_MLOps_Assignment_2$ docker run --rm -p 8000:8000 fastapi_mlopsassignment2:latest
[FO: Started server process [1]
[FO: Waiting for application startup.
[FO: Application startup complete.
[FO: Uvicorn running on http://0.0.0.0:8000 (Press CTRL+C to quit)
[FO: 172.17.0.1:64418 - "GET / HTTP/1.1" 200 OK
[FO: 172.17.0.1:64430 - "GET /favicon.ico HTTP/1.1" 404 Not Found
[FO: 172.17.0.1:55030 - "GET /docs HTTP/1.1" 200 OK
[FO: 172.17.0.1:55030 - "GET /openapi.json HTTP/1.1" 200 OK
[FO: 172.17.0.1:59432 - "GET /api/health HTTP/1.1" 200 OK
[222 10:59:35.177646291 NNPACK.cpp:56] Could not initialize NNPACK! Reason: Unsupported hardware.
[FO: 172.17.0.1:57196 - "POST /api/predict HTTP/1.1" 200 OK
[INFO: Shutting down
[FO: Waiting for application shutdown.
[FO: Application shutdown complete.
[FO: Finished server process [1]
mlopsAssignment2Env) niteshkumar@cseniteshkum...:~/bitsMtech_MLOps_Assignment_2$ docker build -t fastapi_mlopsassignment2 .
[!] Building 1.9s (11/11) FINISHED

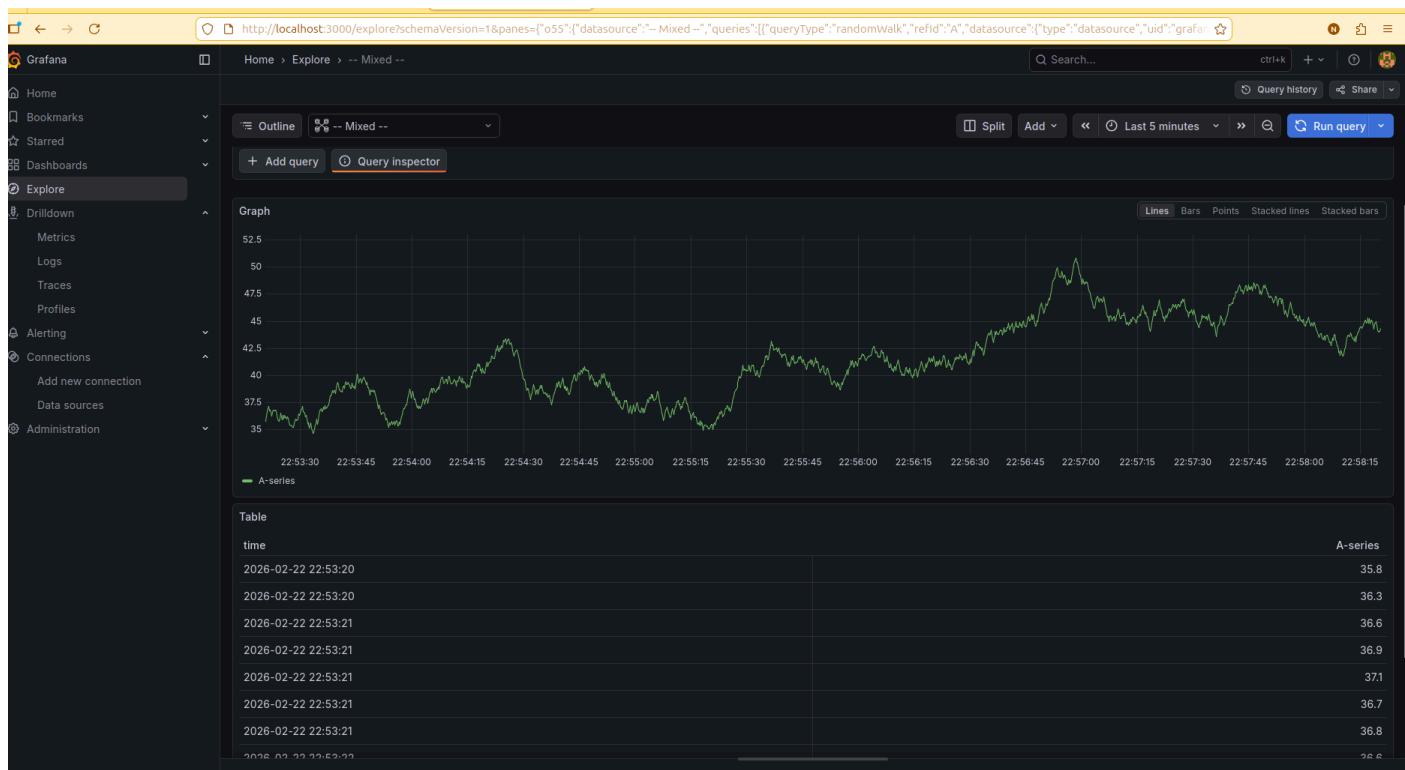
```

## 8. Grafana Tracker

The screenshot shows the Grafana interface with the following details:

- Left Sidebar:** Shows navigation links for Home, Bookmarks, Starred, Dashboards, Explore (selected), Drilldown, Metrics, Logs, Traces, Profiles, Alerting, Connections (Add new connection), Data sources, and Administration.
- Top Bar:** Includes a back/forward button, a search bar with placeholder "http://localhost:3000/explore?schemaVersion=1&panes=[{"o55","datasource":"ef0trsju4n4c","queries":[{"refid":"A","expr":"","range":true,"instant":true,"datasource":{"type":"prometheus"}]}]", a "Run query" button, and other UI elements.
- Central Area:**
  - Prometheus Tab:** Shows a query editor for the "prometheus" data source. It includes sections for "Outline", "prometheus", "Go queryless", "Operations", "Add query", and "Query inspector".
  - Stats Tab:** Displays performance statistics:
 

Total request time	2 ms
Number of queries	0
Total number rows	0



## 9. Repository Structure

The screenshot shows a GitHub code editor interface. The left sidebar displays a project structure with several files and folders under 'src'. The right panel shows the content of 'dataValidation.py'. The code is a Python script for data validation, including imports for sys, os, pathlib, pandas, and numpy, and a function 'validate\_heart\_data' that performs logical checks on a DataFrame.

```

1  """Data validation.
2  """
3  import sys
4  import os
5
6  import pathlib
7
8  sys.path.append(os.path.abspath(os.path.join(os.path.dirname(__file__), '..')))
9
10 import pandas as pd
11 import numpy as np
12
13 from src.dataRead import load_dataset
14 from src.dataClean import clean_data
15
16
17 def validate_heart_data(df):
18     """
19         Performs logical checks to ensure data integrity.
20         Returns True if valid, raises an error or prints warnings otherwise.
21     """
22     validation_passed = True
23
24     # Check 1: No missing values
25     if df.isnull().values.any():
26         print("Validation Failed: Missing values detected.")
27         validation_passed = False
28
29     # Check 2: Logical ranges for medical vitals
30     # (e.g., Blood pressure 'restbps' usually between 80 and 200)
31     if not df['restbps'].between(50, 250).all():
32         print("Validation Warning: Outlier detected in Blood Pressure.")
33
34

```

## Data Versioning (DVC)

The screenshot shows the VS Code interface with the 'EXPLORER' view open. It lists a large number of files and folders, many of which are marked with a small orange circle indicating they are tracked by DVC. The 'DataFiles.dvc' file is selected in the Explorer, and its content is displayed in the main editor area. The content of 'DataFiles.dvc' is a JSON-like configuration for a DVC pointer to 'DataFiles'.

```

1  outs: You, yesterday + Add DVC pointer for DataFiles ...
2  - md5: 855d528753b8d0d9da88003a015da934.dir
3  size: 1713465767
4  nfiles: 37625
5  hash: md5
6  path: DataFiles
7

```

## Requirements\_Pinned.txt

EXPLORER

BITSMTECH\_MLOPS\_ASSIGNMENT\_2

- .dvc
- .gitignore
- config
- .github
- .ipynb\_checkpoints
- .mlopsAssignment2Env
- .pytest\_cache
- .vscode
- app
- artifacts
  - cnn\_model\_full.pt
  - cnn\_model.pt
- Back\_Code\_Files
- Datafiles
- grafana
- htmlcov
- k8s
  - deployment.yaml
  - service.yaml
- logs
- mruns
- temp\_uploads
- tests
- .coverage
- .dockeringore
  - .gitignore
- .coverage.xml
- dataExploration.ipynb
- ! Datafiles.dvc
- docker-compose.yml
- Dockerfile
- milflow.db
- modeTraining.ipynb
- ! prometheus.yml
- pytest.ini
- README.md
- requirements\_fastapi.txt
- requirements\_pinned.txt
- requirements.txt

> OUTLINE

> TIMELINE

> DVC TRACKED

requirements\_pinned.txt

```
1 a2wsgl==1.10.10
2 about-time==4.2.1
3 absl-py==2.4.0
4 addic==2.4.0
5 aiohttpcore==2.26.0
6 aiohappy eyeballs==2.13.3
7 aiohttp==3.13.3
8 aiohttp-asgi-connector==1.1.2
9 aiohttp-retry==2.9.1
10 aioitertools==0.13.0
11 aiosignal==1.4.0
12 aiosqlite==0.22.1
13 alabaster==1.0.0
14 albucore==0.0.24
15 albumen tations==2.0.8
16 alembic==1.18.4
17 alive-progress==3.3.0
18 amqp==5.3.1
19 annotated-doc==0.0.4
20 annotated-types==0.7.0
21 antlr4-python3-runtime==4.9.3
22 anyio==4.12.1
23 appdirs==1.4.4
24 argon2-cffi==25.1.0
25 argon2-cffi-bindings==25.1.0
26 arrow==1.4.0
27 aspiref==3.11.1
28 astroid==4.0.4
29 asttokens==3.8.1
30 astunparse==1.6.3
31 asyncio-lru==2.2.0
32 asynccsh==2.22.0
33 atpublic==7.0.0
34 attrs==25.4.0
35 autograd==1.0
36 autopkg==2.1.2
37 ax-platforme==1.2.3
38 babel==2.18.0
39 beautifulsoup4==4.14.3
40 bentoml==1.4.35
41 billiard==4.2.4
42 black==26.1.0
43 bleach==6.3.0
44 blinker==1.9.0
45 bokeh==3.8.2
46 boto==1.41.5
47 botocore==1.41.5
48 botorch==0.17.0
49 cachetools==7.0.1
50 catboost==1.2.10
51 catfram==23.1.2
```

EXPLORER

BITSMTECH\_MLOPS\_ASSIGNMENT\_2

- .dvc
- .gitignore
- config
- .github
- .ipynb\_checkpoints
- .mlopsAssignment2Env
- .pytest\_cache
- .vscode
- app
- artifacts
  - cnn\_model\_full.pt
  - cnn\_model.pt
- Back\_Code\_Files
- Datafiles
- grafana
- htmlcov
- k8s
  - deployment.yaml
  - service.yaml
- logs
- mruns
- temp\_uploads
- tests
- .coverage
- .dockeringore
  - .gitignore
- .coverage.xml
- dataExploration.ipynb
- ! Datafiles.dvc
- docker-compose.yml
- Dockerfile
- milflow.db
- modeTraining.ipynb
- ! prometheus.yml
- pytest.ini
- README.md
- requirements\_fastapi.txt
- requirements\_pinned.txt
- requirements.txt

> OUTLINE

> TIMELINE

> DVC TRACKED

.gitignore

```
You, 33 minutes ago | 1 author (You)
1 # =====
2 # PYTHON ENVIRONMENT & PACKAGES
3 # =====
4 # Virtual environment directories
5 .mlopsAssignment2Env/
6 .mlopsAssignment2Env/
7 .mlopsAssignment2Env/*
8
9
10 # =====
11 # IDE & EDITOR SPECIFIC
12 # =====
13 # VSCode
14 .vscode/
15 *.code-workspace
16 .vscode-server/
17
18 # =====
19 # Training artifacts / Data Files
20 # =====
21 # *
22 *.jpg
23 *.jpeg
24 *.png
25 *.zip
26
27
28 # =====
29 # JUPYTER NOTEBOOKS
30 # =====
31 # *
32 .ipynb_checkpoints/
33 # *.ipynb
34 desktop.ini
35
36 # =====
37 # LOGGING & OUTPUT
38 # =====
39 /logs/
40 /log/
41 *.log
42 *.log.*
43 nohup.out
44 logs.txt
45
46 # # Output and Results
47 # /output/
48 # /outputs/
```

The file is not displayed in the text editor because it is either binary or uses an unsupported text encoding.

Open Anyway

```
meta.yaml
1 destination_id: m-0377d56c44834102865c663d8ea2efbc
2 destination_type: RUN
3 source_id: m-0377d56c44834102865c663d8ea2efbc
4 source_type: MODEL
5 tags: {}
```

EXPLORER

- BITSMTECH\_MLOPS\_ASSIGNMENT\_2
  - mlruns
  - tags
  - meta.yaml
  - params
  - artifacts
  - models
  - temp\_uploads
  - tests
  - .coverage
  - .dockeringrone
  - .dvcignore
  - .gitignore
  - coverage.xml
  - dataExploration.ipynb
  - Dockerfiles
  - Dockerfile
  - mlflow.db
  - modelTraining.ipynb
  - prometheus.yml
  - pytest.ini
  - README.md
  - requirements\_fastapi.txt
  - requirements\_pinned.txt
  - requirements.txt

! meta.yaml

```

1 conda metadata build recipe - conda's meta.yaml file; https://conda.io/docs/user-guide/tasks/build-packages/define-metadata.html (conda-meta.json) You, yesterday | 1 author (You)
2 artifact_location: file:///media/niteshkumar/SSD_Store_0_nvme/allPythoncodesWithPipEnv/BitsLearning/MLOps_Assignment_2/bitsMtech_MLOps_
3 creation_time: 1771685561004
4 experiment_id: '516697590595924749'
5 lifecycle_stage: active
6 name: CNN_Model_Training_Experiment
7 workspace: default
8

```

artifacts Aa ab No results

EXPLORER

- BITSMTECH\_MLOPS\_ASSIGNMENT\_2
  - .dvc
  - cache
  - tmp
  - .gitignore
  - config
  - ipynb\_checkpoints
  - .mlopsAssignment2Env
  - .pytest\_cache
  - .vscode
  - app
  - artifacts
    - cnn\_model\_full.pt
    - cnn\_model.pt
  - Back\_Code\_Files
  - Datafiles
  - grafana
  - htmlcov
  - k8s
  - logs
  - mlruns
  - temp\_uploads
  - tests
  - .coverage
  - .dockeringrone
  - .dvcignore
  - .gitignore
  - coverage.xml
  - dataExploration.ipynb
  - Dockerfiles
  - Dockerfile
  - mlflow.db
  - modelTraining.ipynb
  - prometheus.yml
  - pytest.ini
  - README.md
  - requirements\_fastapi.txt
  - requirements\_pinned.txt
  - requirements.txt

! README.md

```

1 You, 9 hours ago | author (You)
2
3 . NITESH KUMAR ----- 2024AA05143 ----- 100%
4 . LAWLESH KUMAR ----- 2024AA05149 ----- 100%
5 . VAIBHAV SAREEN ----- 2024AA05023 ----- 100%
6 . VIVEK TRivedi ----- 2024AA05922 ----- 100%
7 . BINDU MANOJ ----- 2024AA05979 ----- 100%
8
9 You, 9 hours ago + Docker and Readme
10 # MLOps Assignment 2 - Cat vs Dog Image Classifier API
11
12 A production-ready machine learning inference service for classifying cat and dog images, built with FastAPI and PyTorch. The project includes comprehensive MLOps practices including containerization, monitoring, testing, and CI/CD readiness.
13
14 ## 🌐 Overview
15
16 This project implements a complete MLOps pipeline for image classification:
17
18 - **Machine Learning**: CNN model trained on cat vs dog dataset using PyTorch
19 - **API Service**: FastAPI-based REST API for real-time inference
20 - **Monitoring**: Prometheus metrics collection and Grafana visualization
21 - **Containerization**: Docker and Docker Compose for easy deployment
22 - **Testing**: Comprehensive unit and integration tests with pytest
23 - **Experiment Tracking**: MLflow for model versioning and experiment tracking
24 - **Data Versioning**: DVC (Data Version Control) for dataset management
25
26
27
28
29 ## 🏠 Architecture
30
31 ...
32
33
34 Client (Browser) --> FastAPI Service --> PyTorch Model
35
36
37
38
39
40
41
42
43
44
45
46
47

```