

CSS 426 ML Ops

Lab - 2

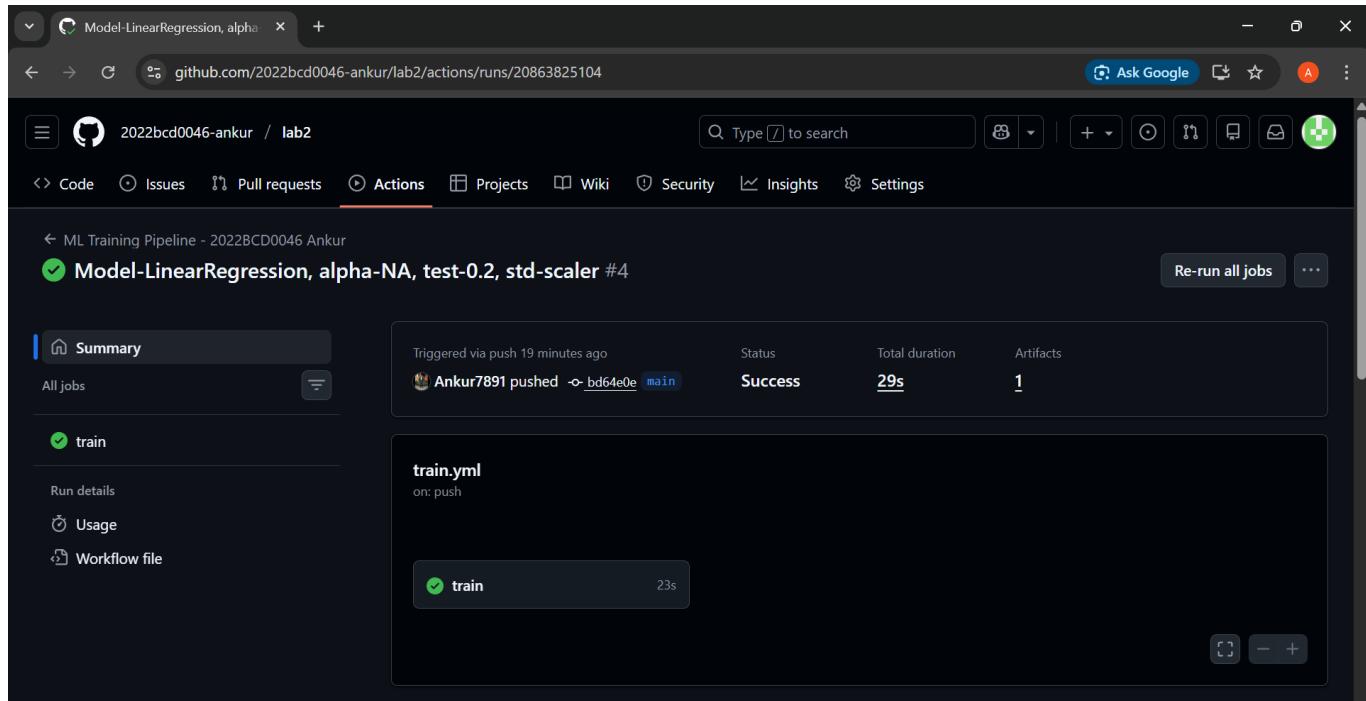
- Name: Ankur Majumdar
 - Roll No. 2022BCD0046
-

1. GitHub Repository Link

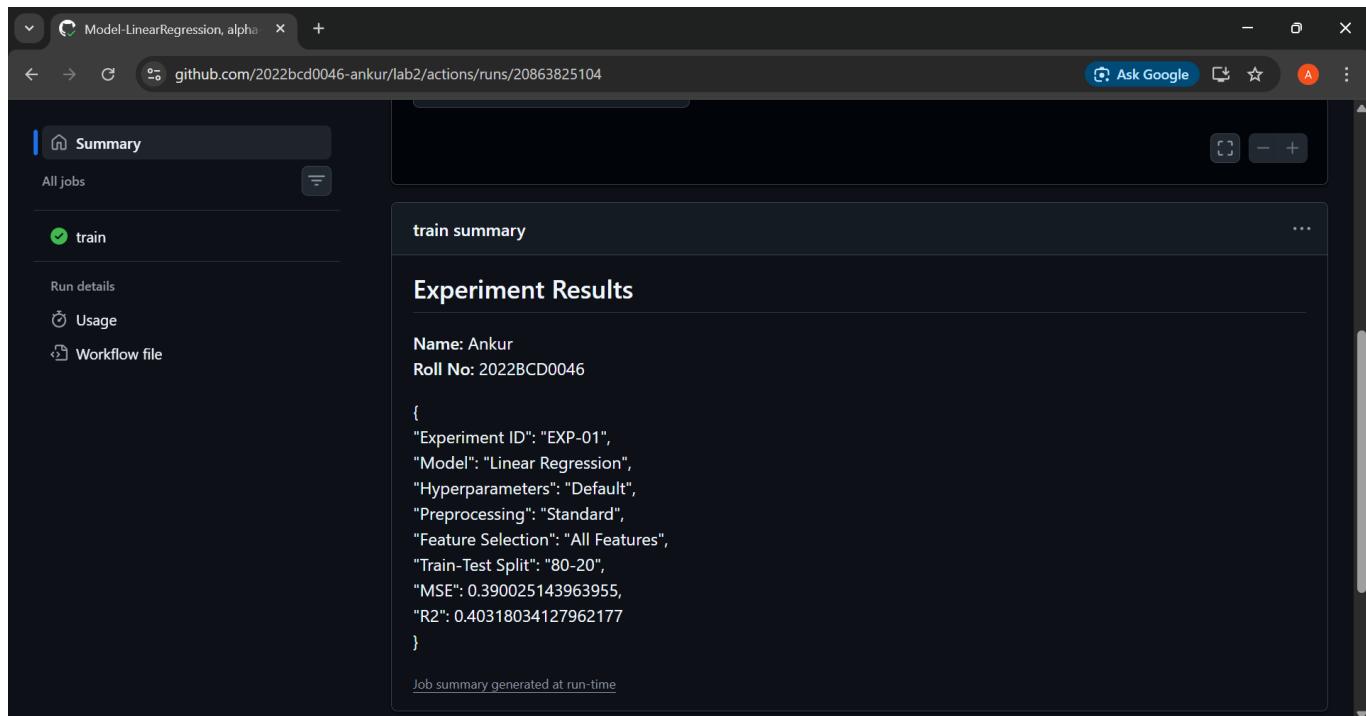
- github.com/2022BCD0046-ankur/lab2

2. Screenshots of Job Summary and Artifacts

- EXP-01: [Run Link](#)



The screenshot shows the GitHub Actions interface for a repository named 'Model-LinearRegression, alpha'. The 'Actions' tab is selected, and the run #4 for the 'Model-LinearRegression, alpha-NA, test-0.2, std-scaler' workflow is displayed. The 'train' job is shown as successful, triggered via push 19 minutes ago. The total duration was 29s and one artifact was produced. The workflow file 'train.yml' is also visible.



The screenshot shows the 'train summary' results for the same GitHub Actions run. It displays the experiment results, including the name (Ankur), roll number (2022BCD0046), and various experimental parameters and metrics. The results are presented in a JSON-like format:

```
{  
  "Experiment ID": "EXP-01",  
  "Model": "Linear Regression",  
  "Hyperparameters": "Default",  
  "Preprocessing": "Standard",  
  "Feature Selection": "All Features",  
  "Train-Test Split": "80-20",  
  "MSE": 0.390025143963955,  
  "R2": 0.40318034127962177  
}
```

A note at the bottom indicates that the job summary was generated at run-time.

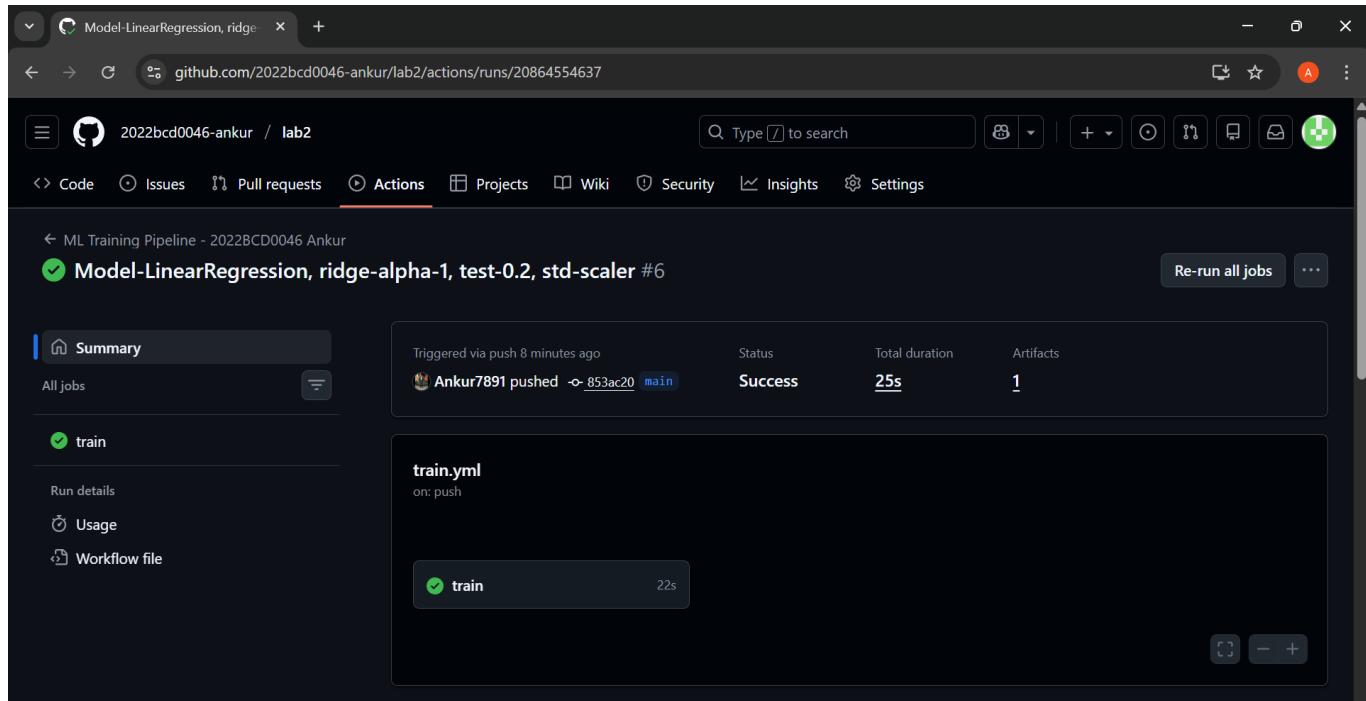
The screenshot shows a GitHub Actions run summary for a Linear Regression model. The top section displays the job configuration with a JSON dump:

```
{  
  "Experiment ID": "EXP-01",  
  "Model": "Linear Regression",  
  "Hyperparameters": "Default",  
  "Preprocessing": "Standard",  
  "Feature Selection": "All Features",  
  "Train-Test Split": "80-20",  
  "MSE": 0.390025143963955,  
  "R2": 0.40318034127962177  
}
```

Below this, a note states "Job summary generated at run-time". The bottom section lists the artifacts produced during runtime:

Name	Size	Digest
experiment-artifacts	1.17 KB	sha256:39fe9ae0980faaddb1d1b44ce85e0...

- EXP-02: [Run Link](#)



The screenshot shows the GitHub Actions interface for a repository named 'Model-LinearRegression, ridge-alpha-1, test-0.2, std-scaler'. A specific run (#6) is selected, triggered via push 8 minutes ago by user Ankur7891. The status is Success, total duration is 25s, and there is 1 artifact. The 'train' job is highlighted with a green checkmark. The workflow file 'train.yml' shows a single step 'train' with a duration of 22s. A 'train summary' section displays experiment results: Name: Ankur, Roll No: 2022BCD0046, and a JSON object detailing the experimental setup.

Triggered via push 8 minutes ago

Ankur7891 pushed -o 853ac20 main

Status: Success, Total duration: 25s, Artifacts: 1

train

train.yml

on: push

train

22s

train summary

Experiment Results

Name: Ankur
Roll No: 2022BCD0046

```
{  
  "Experiment ID": "EXP-02",  
  "Model": "Linear Regression",  
  "Hyperparameters": "Ridge Alpha-1",  
  "Preprocessing": "Standard",  
  "Feature Selection": "All Features",  
  "Train-Test Split": "80-20",  
  "MSE": 0.3900365808244968,  
  "R2": 0.4031628405022718  
}
```

Job summary generated at run-time

The screenshot shows a GitHub Actions run summary for a Linear Regression model. The URL is github.com/2022bcd0046-ankur/lab2/actions/runs/20864554637. The summary includes the following details:

Roll No: 2022BCD0046

```
{
  "Experiment ID": "EXP-02",
  "Model": "Linear Regression",
  "Hyperparameters": "Ridge Alpha-1",
  "Preprocessing": "Standard",
  "Feature Selection": "All Features",
  "Train-Test Split": "80-20",
  "MSE": 0.3900365808244968,
  "R2": 0.4031628405022718
}
```

Job summary generated at run-time

Artifacts
Produced during runtime

Name	Size	Digest
experiment-artifacts	1.07 KB	sha256:01e48ad3687123413dc766d17281a...

- EXP-03: [Run Link](#)

The image shows two screenshots of a GitHub Actions run summary for a repository named "Model-Decision Tree, random-state-42, test-0.2, std-scaler, corr-feat...".

Screenshot 1: Run Summary

This screenshot displays the main run summary page. It includes a summary card for the entire run and a detailed view of the "train" job.

- Summary Card:** Triggered via push 2 minutes ago by Ankur7891 pushed -o 25caa3d main. Status: Success. Total duration: 30s. Artifacts: 1.
- Job Details:** train (train.yml) triggered on push. Duration: 28s. Status: Success.

Screenshot 2: Experiment Results

This screenshot shows the "Experiment Results" section of the run summary. It provides a summary of the experiment and its configuration.

Experiment Summary: train summary

Experiment Results:

Name: Ankur
Roll No: 2022BCD0046

```
{  
  "Experiment ID": "EXP-03",  
  "Model": "Decision Tree",  
  "Hyperparameters": "Random State-42",  
  "Preprocessing": "Standard",  
  "Feature Selection": "Correlation-based",  
  "Train-Test Split": "80-20",  
  "MSE": 0.721875,  
  "R2": -0.10461901702057697  
}
```

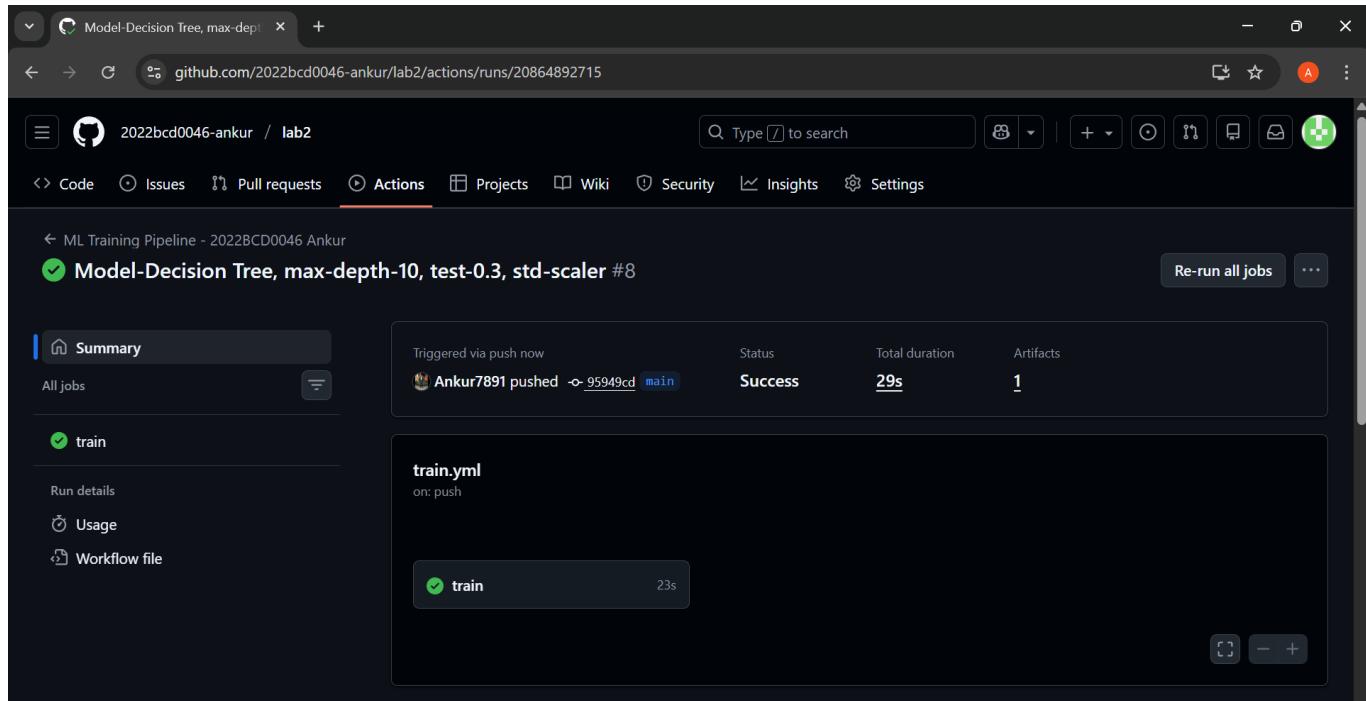
Job summary generated at run-time

The screenshot shows a GitHub Actions run summary for a job named "train". The job was run by Ankur with Roll No: 2022BCD0046. The configuration parameters for the experiment are listed as follows:

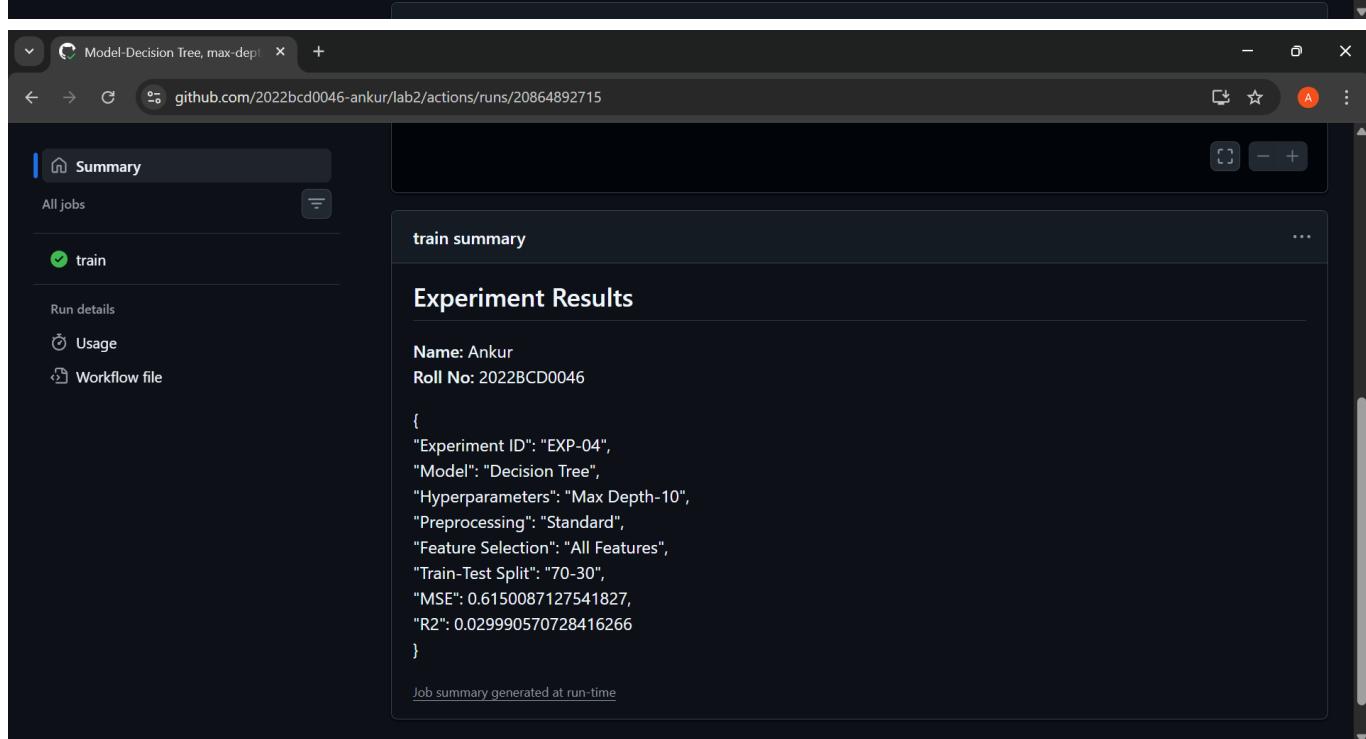
```
Name: Ankur  
Roll No: 2022BCD0046  
{  
  "Experiment ID": "EXP-03",  
  "Model": "Decision Tree",  
  "Hyperparameters": "Random State-42",  
  "Preprocessing": "Standard",  
  "Feature Selection": "Correlation-based",  
  "Train-Test Split": "80-20",  
  "MSE": 0.721875,  
  "R2": -0.10461901702057697  
}  
  
Job summary generated at run-time
```

The "Artifacts" section shows a single artifact named "experiment-artifacts" which is 10 KB in size and has a SHA-256 digest of sha256:9a5d398bd3fb79e956860beea7b65e... . There are download and delete icons next to the artifact entry.

- EXP-04: [Run Link](#)



The screenshot shows the GitHub Actions interface for a repository named 'Model-Decision Tree, max-depth-10'. The 'Actions' tab is selected. A green checkmark indicates a successful run labeled 'Model-Decision Tree, max-depth-10, test-0.3, std-scaler #8'. The 'Summary' card shows the job was triggered via a push to the 'main' branch by user 'Ankur7891'. The status is 'Success' with a duration of '29s' and one artifact. The 'train' job is detailed in a separate card, showing it was triggered on a push and completed in 23 seconds.



The screenshot shows the GitHub Actions interface for the same repository. The 'train' job is selected. The 'Experiment Results' section displays the following details:

```
Name: Ankur  
Roll No: 2022BCD0046  
{  
  "Experiment ID": "EXP-04",  
  "Model": "Decision Tree",  
  "Hyperparameters": "Max Depth-10",  
  "Preprocessing": "Standard",  
  "Feature Selection": "All Features",  
  "Train-Test Split": "70-30",  
  "MSE": 0.6150087127541827,  
  "R2": 0.029990570728416266  
}
```

A note at the bottom states: 'Job summary generated at run-time'.

The screenshot shows a GitHub Actions job summary for a run named "Ankur" with Roll No: 2022BCD0046. The job details include the experiment ID ("EXP-04"), model type ("Decision Tree"), hyperparameters ("Max Depth-10"), preprocessing ("Standard"), feature selection ("All Features"), train-test split ("70-30"), MSE (0.6150087127541827), and R2 (0.029990570728416266). Below the job summary is a section for artifacts, showing a single artifact named "experiment-artifacts" with a size of 6.84 KB and a SHA-256 digest. The artifact has download and delete icons.

- Overall Runs...

The screenshot shows the GitHub Actions workflow runs page for the repository "ML Training Pipeline - 2022BCD0046 Ankur". The sidebar on the left shows management options like Caches, Attestations, Runners, Usage metrics, and Performance metrics. The main area displays four workflow runs, each with a green checkmark indicating success. The first run is for "Model-Decision Tree, max-depth-10, test-0.3, std-sca..." pushed by Ankur7891. The second run is for "Model-Decision Tree, random-state-42, test-0.2, std-..." pushed by Ankur7891. The third run is for "Model-LinearRegression, ridge-alpha-1, test-0.2, std-..." pushed by Ankur7891. The fourth run is for "Model-LinearRegression, alpha-NA, test-0.2, std-scal..." pushed by Ankur7891. Each run entry includes a timestamp, status (main), and three-dot more options menu.

3. Analysis Questions

1. How did GitHub Actions improve reproducibility?
 - GitHub Actions ensured that every experiment ran in a fixed, automated environment with the same dependency versions and execution steps, making results reproducible across runs.
2. How easy was it to compare results across runs?
 - Comparison was easy because each run produced metrics in the Job Summary and downloadable JSON artifacts, allowing side-by-side inspection.

3. What role does Git commit history play in experiment tracking?

- Each commit represents a unique experiment configuration, providing a clear and traceable history of changes and their corresponding results.

4. Benefits compared to Lab 1

- Unlike Lab 1, experiments were automated, reproducible and centrally logged, reducing manual effort and error.

5. Limitations

- Hyperparameter changes are still manual.