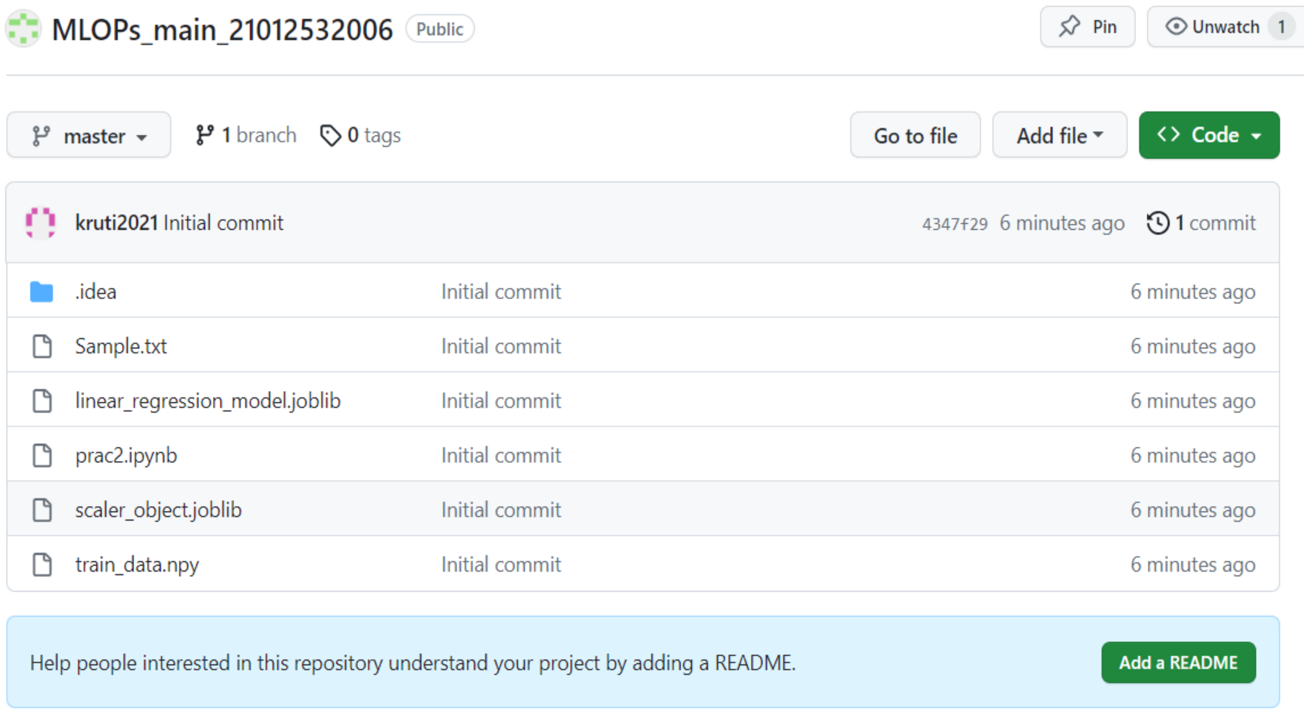
# Practical-3

# Generation of Reproducible and Interactive ML Project.

**Task 1:** Create the Github repository for the house rate prediction project created in practical 2.



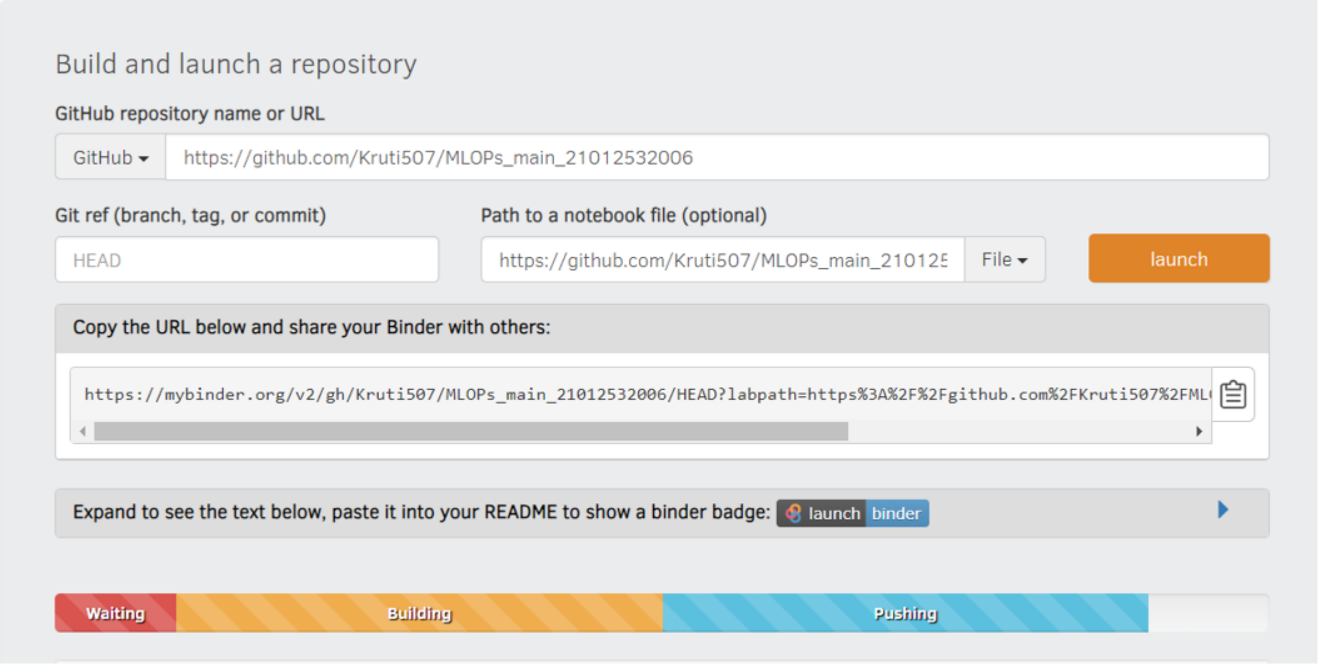
MLOPs_main_21012532006 Public

kruti2021 Initial commit 4347f29 6 minutes ago 1 commit

| | | |
|---|---|---|
| .idea | Initial commit | 6 minutes ago |
| Sample.txt | Initial commit | 6 minutes ago |
| linear_regression_model.joblib | Initial commit | 6 minutes ago |
| prac2.ipynb | Initial commit | 6 minutes ago |
| scaler_object.joblib | Initial commit | 6 minutes ago |
| train_data.npy | Initial commit | 6 minutes ago |

Help people interested in this repository understand your project by adding a README.

**Task 2:** Integrate your repository with the binder to make your project interactive. (Hint: refer to the following link for the steps: ([https://mybinder.org/](https://mybinder.org/))



Build and launch a repository

GitHub repository name or URL

GitHub https://github.com/Kruti507/MLOPs_main_21012532006

Git ref (branch, tag, or commit)

HEAD

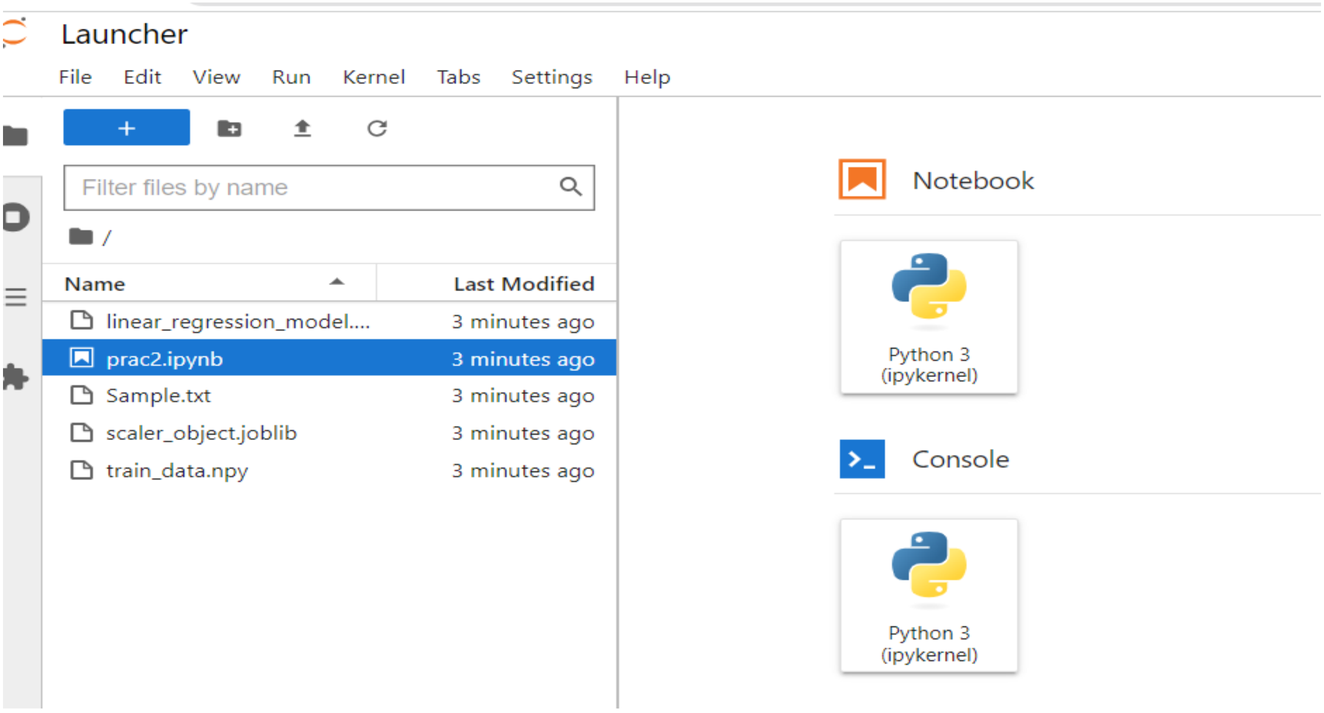
Path to a notebook file (optional)

https://github.com/Kruti507/MLOPs_main_21012Ɛ File

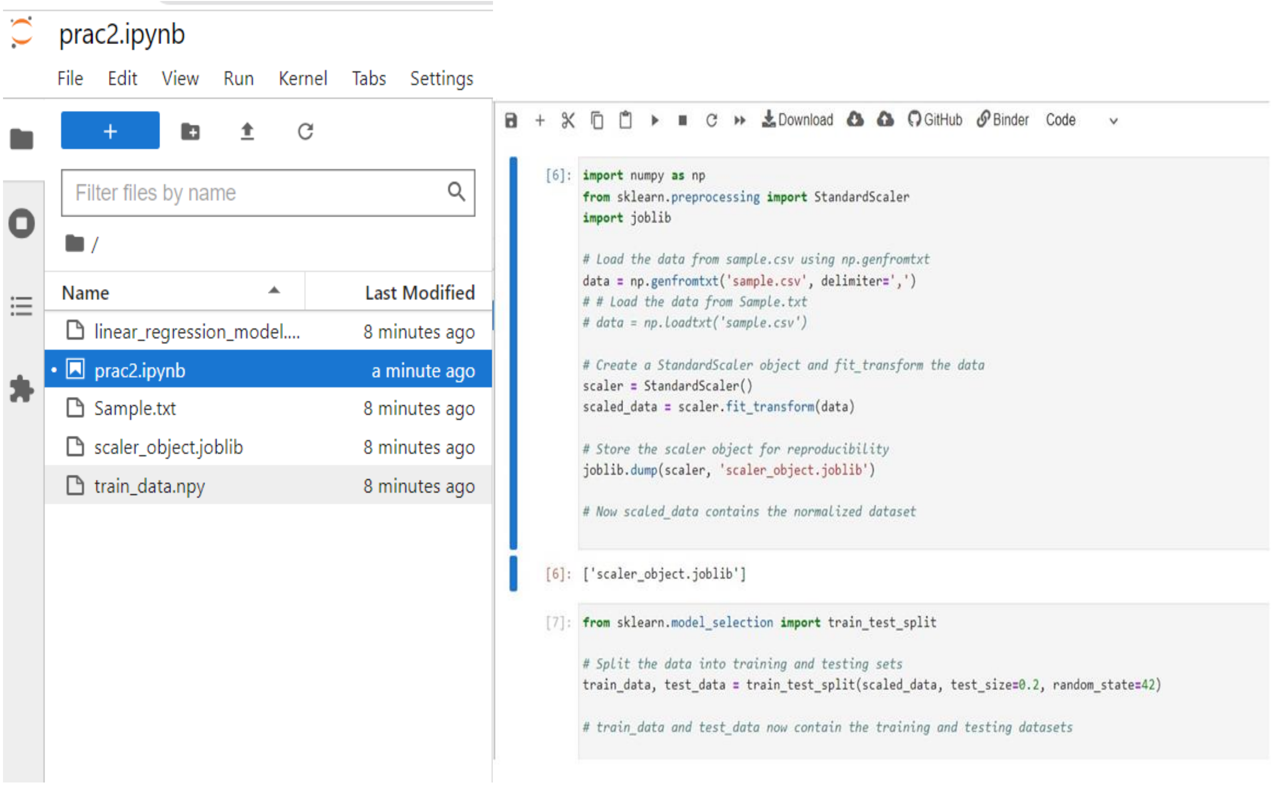
Copy the URL below and share your Binder with others:

```
https://mybinder.org/v2/gh/Kruti507/MLOPs_main_21012532006/HEAD?labpath=https%3A%2F%2Fgithub.com%2FKruti507%2FML
```

Expand to see the text below, paste it into your README to show a binder badge:



| Name | Last Modified |
|---|---|
| linear_regression_model.... | 3 minutes ago |
| prac2.ipynb | 3 minutes ago |
| Sample.txt | 3 minutes ago |
| scaler_object.joblib | 3 minutes ago |
| train_data.npy | 3 minutes ago |



| Name | Last Modified |
|---|---|
| linear_regression_model.... | 8 minutes ago |
| prac2.ipynb | a minute ago |
| Sample.txt | 8 minutes ago |
| scaler_object.joblib | 8 minutes ago |
| train_data.npy | 8 minutes ago |

[6]:
```
import numpy as np
from sklearn.preprocessing import StandardScaler
import joblib

# Load the data from sample.csv using np.genfromtxt
data = np.genfromtxt('sample.csv', delimiter=',')
# # Load the data from Sample.txt
# data = np.loadtxt('sample.csv')

# Create a StandardScaler object and fit_transform the data
scaler = StandardScaler()
scaled_data = scaler.fit_transform(data)

# Store the scaler object for reproducibility
joblib.dump(scaler, 'scaler_object.joblib')

# Now scaled_data contains the normalized dataset
```

[6]:
```
['scaler_object.joblib']
```

[7]:
```
from sklearn.model_selection import train_test_split

# Split the data into training and testing sets
train_data, test_data = train_test_split(scaled_data, test_size=0.2, random_state=42)

# train_data and test_data now contain the training and testing datasets
```