Learn ML flow in details.

End-to-end Chest-Cancer-Classification-Using ML flow & DVC

1. Introduction and GitHub Repository Setup
2. Project Template Creation
3. Project Setup and Requirements Installation
4. Logging, Utils and Exception Module
5. Project Workflows
6. All components notebook experiment
7. Training Pipeline
8. MLflow (MLOps Tool) – for experiments tracking and Model Registration
9. DVC (MLOps tool) – for Pipeline tracking and implementation
10. Prediction pipeline and user app creation
11. Docker
12. Final CI/CD deployment on AWS and Azure

Prerequisite

1. Python programming (OOPs Concepts Required)
2. Deep learning – Object Classification Basic and tensorflow 2.xAPI
3. AWS and Azure Account
4. Your dedication

What is Adenocarcinoma?

Adenocarcinoma is a type of cancer that starts in mucus-producing(glandular) cells. Many organs have these types of cells and adenocarcinoma can develop in any of these organs.

Sklearn link:

Run on anaconda environment: python MLflow\_train.py #after running it, mlruns folder is created and then will get an experiment with the unique ID.

Inside mlruns folder, we can see meta.yaml file:  
  
 artifact\_uri: file:///D:/Classfication\_Model\_Using\_MLflow/mlruns/0/8388c0cfb99b4f61baecf94e6e67f97f/artifacts

end\_time: 1700899911873

entry\_point\_name: ''

experiment\_id: '0'

lifecycle\_stage: active

run\_id: 8388c0cfb99b4f61baecf94e6e67f97f

run\_name: redolent-ox-402

run\_uuid: 8388c0cfb99b4f61baecf94e6e67f97f

source\_name: ''

source\_type: 4

source\_version: ''

start\_time: 1700899907998

status: 3

tags: []

user\_id: jnote

Inside conda.yaml file shows which packages are using in this project.

channels:

- conda-forge

dependencies:

- python=3.8.18

- pip<=23.3.1

- pip:

  - mlflow<3,>=2.2

  - cloudpickle==2.2.1

  - numpy==1.24.4

  - scikit-learn==1.3.2

  - scipy==1.10.1

name: mlflow-env

ML Model configuration:

artifact\_path: model

flavors:

  python\_function:

    env:

      conda: conda.yaml

      virtualenv: python\_env.yaml

    loader\_module: mlflow.sklearn

    model\_path: model.pkl

    predict\_fn: predict

    python\_version: 3.8.18

  sklearn:

    code: null

    pickled\_model: model.pkl

    serialization\_format: cloudpickle

    sklearn\_version: 1.3.2

mlflow\_version: 2.2.2

model\_uuid: af2bfbf221704c389ee1354222b973bb

run\_id: 8388c0cfb99b4f61baecf94e6e67f97f

utc\_time\_created: '2023-11-25 08:11:48.092594'

To display running times, we can go through ML flow ui.

How to set up ML Flow in remote server (dagshub.com)? #What is dagshub?

To run in ML flow ui: mlflow ui