DVC (Data Vernion Control)

L. basics

L. where applicable

L. projecti — basics

L. Aklear ML

L. Tensorflow (Beeplearing)

L. Pytorch

Ly Data Versioning

Ly model Versioning artifacts.

→ Automate MI Pipelines with DVC

→ Stage OI load & Save py

→ Stage O2 Split data py.

→ Stage O3 train py

→ Stage O4 evaluate py

Mersion Controlled Copies,

Chauges,

Chauges,

Integrates with Git

It build pipelines.

1 + Eloge of 1

wife and only interested to

Poreprinder

And it a the set in

DVC (Data Vernion Control)

L. basics

L. where applicable

L. projects -> basics

L. Aklear ML

L. Tensorflow (Deeplearing)

L. Pytorch

Ho This is helpful for Experiment toacking

Ly Data Versioning

Ly model Versioning artifacts.

→ Automate MI Pipelines with DVC

Ly Stage 01 load & Save py

Ly Stage 02 Split data py.

Ly Stage 03 train py

Ly Stage 03 train py

Ly Stage 04 evaluate py

Mersion Controlled Copies,

Chauges,

Chauges,

Integrates with Git

Scripts & notebooks

It build pipelines.

slip of the state of the first file.

L. Created Stage-01.py file & Stage 02.py file. evaluate Stage 01. Py write read Stage -02 - inside artifacts

artifacts txt tie olp.

Lin most of me pipeline we write something Extresto Create afile. & then read it.

LADAG Brect Acyclic graph should follow here we write Create & then write. This flow should follow read

-> create . dvc , dvc ignore files (we are initializing them) dyc init, git init Ly Create duc. Yami tile

[duc. Yam] \_\_\_, helps to execute (duc. lock in Sequential tile) manner.

1) in duc. Yami we give

Stages HOURS HOURS

→ Stage 01

-> cond (running Stage of file)

La dependencies (files needed donun and

Ind who art made I of the Stage of

Ly Stage 02 eardaging helies in and I condition

>> dependency

1 old's if any.

at the end this creates duc. Lock file

```
L+ it consists of properine of Stages.
             Lim this along with Given variables in duc. Yam 1
                file it has ar generate mds (which is a hash
                value/key) a also size of tile + has.
         Stageoz dvc. Yami - dvc. lock
                                     - dependences
Amonationes) more who ilmor often Lygile Sizes show a
                      in de any of the ter
                  - Try July Sies consider all commonfiles

required tor

utils sies consider all below files.
   Automate ML pipelines with DVC
          Save low cont into to to at tont +1
               Lioutside of Src folder we will have (yout folder
                          Ly Astifacts folder
                              Lit Conmits rawdata, processed
                       model, stores metrics,
                            L+ params yaml
           of present stars with responsible for having all parame-
                            may hamters responsible for training.
                            Ly duc yaml
                                     Lit consult of all stages steps
        pronound iglo improvation of that occurred, we mention
  in all is stiff impossible assured and those here
                                          - Atoge Ol. load & Save . py
                                           -> Stage 02 Aplitdata.py
         from a contract in comp front
                                           -> stage 03 trainipy
                           Ly Config Yaml - mage of Evaluate py
```

wc.lock file

After that we create Requirements txt file.

Ly working on DVC pipeline → Create Src folder

Lyin thu Oreate Utils folders.

L+ in utils create --init\_-. py file

Linside Src fader Create -- init -- py file & that it consider this as package.

Ly Create Params. Yami, Config. Yami, dvc. Yami (wing touch and) invide config dader

enter files consider of all commonfer. Config. Yaml

Hodata Source: <dataset link>

Lyartifacti;

Hartifacti-dir: artifacti Hraw Local-dir: raw-local-dir rate ( hor) and the sub-raw-local-file: data-csv 1- Split-data:

Ly In Src Jolder

Li create load save py file

Lygnide utils folder create utils py file. promod system (am) L+ imput os AN COUNTY DOUTED WITH

Ly ilp path to yam 1 olp: dictionary

ward around by This takes configurems file asilp & reads it & gives olp as dictionary form of content in conting. Yaml

Ly and then we need to areata a Setup.py file.

Lyilp: readme.md file (it reads the file)

Ex Setup(....

Later top-level-txt...

Lithi Netuppy tile & SYC becomes local packages to
TO install local packages, -e. in requirements txt.

Ethen run requirements txt.

Ly Now we need to work on load-save-py file.

Lilp: Org parser()

Config. Yam ((31) params fam) file

Ly olp: name of given config file.

(or) content of the config file.

I ip: path to data source in contig. Yam!

Who olp: It reads the data in path & gives it.

LANOW as we got the data from remote path, we need to save the data that we got from remote path into artifacts directory

Ly Create Im mide Thu load-save py file

in all utils: py.

Holp: Empty list

How Considering Config. Yami file we need to Save datasetin local din for that areating path to directory.

1> 91p: artifacts.dir raw\_local-dir raw-local-file raw-local-dir-path apply life & Six becomes local packages to raw-local-file-path. most fout packages, et. in resurance tot.

Opp directory gets created of your was about no show or harr our and

+Till now we shave finished stage of (loading data)

L-Now adding stage to duc Yami file

of ilpr Stages:

Cond: run load-savery file with Config Yami asilP

dependes > all-utils-py Contig . Yaml dry storm met to me to the load sale py

Olp. data CSV. That has been

L) commit Changes Inceded.

of other golden deliners mother the

Li Now let us go further & add next stage (or) which is splitting of data.

L+ Create file in src folder Src/Stage-02-Spt-data.py.

Eplit ratio, vandom-state

Olp: Odf from data as y from path

L> gn: Save-local-db - Lip: data, data-porth L> olp: - Save file @ adatapath

→ trandata path - ilp: artifacti\_dir, Spirtdata dir, Hename.

La Create - directory - ilp: artifacti dir, Spirtdata dir.

Lonce we do all the Olp: - train.csv

Ly Now we can add thus to alvayam! Itage.

Ly mention deps, params, outs Eall in duc. Yam!

Stage.

→ Save-local of -> ilp: data, datapath

L> defining directory path for storing split data from Config.

La Now Create stage - 03 - training wing touch and in Sic fold und utils folder

La in train we need to Specify directories & paths along with hyperparameters

→ Greate a trained model

→ & Save the model to disk.

ip: In to train model

paths q directories

CSV file

model params

olpstrained model in model-dir

the mention dependance of the contraction of the co

1. Sour local db => 1/p date delapath

Les defining directory hasts for storing spect date from Ording

Ly Now Create another tile eva Stage-03- evaluate-py in 8rc folder.

Ly "Ip: to ained - model
text-data
textypredicted Dalues
a Ghal-Values
reports-dir
Scores
Scores-filename

Olp: predicted values

rmse, mae, rz (metrics)

Scores dir

Scores filename

Scores dir path

Sores-filepath

Lif we want to add a stage where it gives best scores & best model among muttiple models.

