STEPS TO WRITE CODE LINE BY LINE FOR THIS PROJECT OF CHICKEN DISEASE

1. Create a Git repository (add readme, git ignore file as python and license as MIT)
2. Open Git bash, navigate to your cwd and get clone the repo

git clone (paste the repo link)

1. Create a template.py file to setup the file structure and write the following code in it.
2. import os
3. from pathlib import Path  #this will help to recognise that it is windows path and ignore the "/"
4. import logging
5. logging.basicConfig(level=logging.INFO, format='[%(asctime)s]: %(message)s: %(lineno)s')
6. project\_name = "cnnClassifier"
7. list\_of\_files=[
8. ".github/workflows/.gitkeep",       #used when deploying the app, writing all CI/CD commands (gitkeep is used, when if you commit the code and any file is empty)
9. f"src/{project\_name}/\_\_init\_\_.py",  #\_\_init\_\_.py is a constructor file which is needed to use that file methods in another file
10. f"src/{project\_name}/components/\_\_init\_\_.py",
11. f"src/{project\_name}/utils/\_\_init\_\_.py",
12. f"src/{project\_name}/config/\_\_init\_\_.py",
13. f"src/{project\_name}/config/configuration/\_\_init\_\_.py",
14. f"src/{project\_name}/pipeline/\_\_init\_\_.py",
15. f"src/{project\_name}/entity/\_\_init\_\_.py",
16. f"src/{project\_name}/constants/\_\_init\_\_.py",
17. "config/config.yaml",
18. "dvc.yaml",
19. "params.yaml",
20. "requirements.txt",
21. "setup.py",
22. "research/trails.ipynb",
23. "templates/index.html"
24. ]
25. for filepath in list\_of\_files:
26. filepath = Path(filepath)
27. filedir, filename = os.path.split(filepath)
28. if filedir != "":
29. os.makedirs(filedir, exist\_ok=True)
30. logging.info(f"Creating directory: {filedir} for the file: {filename}")
31. if (not os.path.exists(filepath)) or (os.path.getsize(filepath)==0):
32. with open(filepath, "w") as f:
33. pass
34. logging.info(f"Creating empty file: {filepath}")
35. else:
36. logging.info(f"{filename} is already existing")
37. Write the requirements.txt and add the below code
38. tensorflow
39. pandas
40. dvc
41. notebook
42. numpy
43. matplotlib
44. seaborn
45. python-box==6.0.2
46. pyYAML
47. tqdm
48. ensure==1.0.2
49. joblib
50. types-PyYAML
51. scipy
52. FLask
53. Flask-Cors

5. So after writing these required libraries, we need to create our new environment variable for our project, so write the following code in the terminal

conda create –n “name\_for\_variable(chicken” python=3.8 –y

6. Activate the env : conda activate chicken

7. Now install the requirements using the following code-

pip install –r requirements.txt

8. We need to log our actions so for that inside the src/cnnClassifier contructor file made, write the below code

import os

import sys

import logging

logging\_str = "[%(asctime)s: %(levelname)s: %(module)s: %(message)s]"

log\_dir = "logs"

log\_filepath = os.path.join(log\_dir, "running\_logs.log")

os.makedirs(log\_dir, exist\_ok=True)

logging.basicConfig(

    level= logging.INFO,

    format= logging\_str,

    handlers= [

        logging.FileHandler(log\_filepath),          #this will add the dir to the given path

        logging.StreamHandler(sys.stdout)           #tthis wil show the log on your terminal as well

    ]

)

logger = logging.getLogger("cnnClassifierLogger")

This will create a file named logs and will store every logging information inside it.

9. We have to create utils, just to store the repetitive functions used in the projects, so create common.py inside the utils folder

\* ensure\_annotations – This is used to indicate error such as when int tries to multiply with str and which python allows normally but using this decorator, we can debug the code and understand the issue.

\* Config\_box – using this we can access the keys of dictionaries using “.” Operator, we have import Box and then write while defining a dictionary = ConfigBox({‘key’:’value’})

The below code is needed to put in common.py folder

import os

from box.exceptions import BoxValueError

import yaml

from cnnClassifier import logger

import json

import joblib

from ensure import ensure\_annotations

from box import ConfigBox

from pathlib import Path

from typing import Any

import base64

@ensure\_annotations

def read\_yaml(path\_to\_yaml: Path) -> ConfigBox:

    """read yaml file and returns

    Args:

        path\_to\_yaml(str): path like input

    Raises:

        ValueError: if yaml file is empty

        e: empty file

    Returns:

        ConfigBox: ConfigBox type

    """

    try:

        with open(path\_to\_yaml) as yaml\_file:

            content =  yaml.safe\_load(yaml\_file)

            logger.info(f"yaml file: {path\_to\_yaml} loaded successfully")

            return ConfigBox(content)

    except BoxValueError:

        raise ValueError("yaml is empty")

    except Exception as e:

        raise e

@ensure\_annotations

def create\_directories(path\_to\_directories: list, verbose=True):            #this is used to make the directories in the artifacts folder i.e.data ingestion, model, etc

    """"creaate list of directories

    Args:

        path\_to\_directories (list): list of path of directories

        ignore\_log (bool, optional): ignore if multiple dirs is to be created. Defaults to False

    """

    for path in path\_to\_directories:

        os.makedirs(path, exist\_ok=True)

        if verbose:

            logger.infor(f"Created directory at: {path}")

def save\_json(path: Path, data:dict):           #this json file is created to save the accuracy and loss after evalaution model.

    """ save json data

    Args:

        path (Path): path to json file

        data (dict): data to be saved in json file.

    """

    with open(path, "w") as f:

        json.dump(data, f, indent=4)

    logger.info(f"json file saved at {path}")

@ensure\_annotations

def load\_json(path: Path) -> ConfigBox:

    """ load json files data

    Args:

        path (Path): path to json file

    Returns:

        ConfigBox: data as class attributes instead of dict

    """

    with open(path) as f:

        content=json.load(f)

    logger.info(f"json file loaded successfully from {path}")

    return ConfigBox(content)

@ensure\_annotations

def save\_bin(data: Any, path: Path) -> ConfigBox:

    """ save binary file

    Args:

        data (Any): data to be saved as binary

        path (Path): path to binary file

    """

    joblib.dump(value=data, filename=path)

    logger.info(f"binary file saved at {path}")

@ensure\_annotations

def load\_bin(path: Path) -> ConfigBox:

    """ load binary data

    Args:

        path (Path): path to binary file

    Returns:

        data (Any): object stored in the file

    """

    data = joblib.load(path)

    logger.info(f"binary file loaded from {path}")

    return data

@ensure\_annotations

def get\_size(path: Path) -> ConfigBox:

    """ get size in Kb

    Args:

        path (Path): path of the file

    Returns:

        str: size in Kb

    """

    size\_in\_kb = round(os.path.getsize(path)/1024)

    return f"~ {size\_in\_kb} KB"

def decodeImage (imgstring, filename):

    imgdata = base64.b64decode(imgstring)

    with open(filename, 'wb') as f:

        f.write(imgdata)

        f.close()

def enodeImageIntoBase64(croppedImagePath):

    with open(croppedImagePath, "rb") as f:

        return base64.b64dencode(f.read())

10. Now we have discussed the workflow, that which file is needed to start with

##Workflows

1. Update config.yaml

2. Update secrets.yaml [optional]

3. update params.yaml

4. update the entity

5. update the configuration manager in src config

6. Update the components

7. Update the pipeline

8. Update the main.py

9. update the dvc.yaml

11. After that we drag and drop the zip dataset file in github

12. After this we have to prepare the data ingestion notebook and write the code there, and once when the code perfect we have to put it in modular coding format

13. Inside confgi.yaml, we need to write the artifacts folder and it’s subfolders name because even if you want to change any folder names, you change it from here.

artifacts\_root: artifacts

data\_ingestion:

  root\_dir: artifacts/data\_ingestion

  source\_URL: https://github.com/uv0037/Chicken\_disease\_classification/raw/main/Chicken-fecal-images.zip

  local\_data\_file: artifacts/data\_ingestion/data.zip

  unzip\_dir: artifacts/data\_ingestion

14. Now to write the code in data ingestion file.

from dataclasses import dataclass

from pathlib import Path

@dataclass(frozen=True)

class DataIngestionConfig:

    root\_dir: Path

    source\_URL: str

    local\_data\_file: Path

    unzip\_dir: Path

from cnnClassifier.constants import \*

from cnnClassifier.utils.common import read\_yaml, create\_directories

class ConfigurationManager:

    def \_\_init\_\_(

        self,

        config\_filepath = CONFIG\_FILE\_PATH,

        params\_filepath = PARAMS\_FILE\_PATH):

        self.config = read\_yaml(config\_filepath)

        self.params = read\_yaml(PARAMS\_FILE\_PATH)

        create\_directories([self.config.artifacts\_root])

    def get\_data\_ingestion\_config(self) -> DataIngestionConfig:

        config = self.config.data\_ingestion

        create\_directories([config.root\_dir])

        data\_ingestion\_config = DataIngestionConfig(

            root\_dir = config.root\_dir,

            source\_URL = config.source\_URL,

            local\_data\_file = config.local\_data\_file,

            unzip\_dir = config.unzip\_dir

        )

        return data\_ingestion\_config

import os

import urllib.request as request

import zipfile

from cnnClassifier import logger

from cnnClassifier.utils.common import get\_size

class DataIngestion:

    def \_\_init\_\_(self, config: DataIngestionConfig):

        self.config=config

    def download\_file(self):

        if not os.path.exists(self.config.local\_data\_file):

            filename, headers = request.urlretrieve(

                url = self.config.source\_URL,

                filename = self.config.local\_data\_file

            )

            logger.info(f"{filename} download with the below headers \n{headers}")

        else:

            logger.info(f"File already exists of size: {get\_size(Path(self.config.local\_data\_file))}")

    def extract\_zip\_file(self):

        unzip\_path = self.config.unzip\_dir

        os.makedirs(unzip\_path, exist\_ok=True)

        with zipfile.ZipFile(self.config.local\_data\_file, 'r') as zip\_ref:

            zip\_ref.extractall(unzip\_path)

try:

    config = ConfigurationManager()

    data\_ingestion\_config = config.get\_data\_ingestion\_config()

    data\_ingestion = DataIngestion(config=data\_ingestion\_config)

    data\_ingestion.download\_file()

    data\_ingestion.extract\_zip\_file()

except Exception as e:

    raise e

The above code will create the folder artifacts, download the dataset zip file from github link and then unzip it.

15. Create a new file in entity as config\_entity.py and paste the following code  
  
from dataclasses import dataclass

from pathlib import Path

@dataclass(frozen=True)

class DataIngestionConfig:

    root\_dir: Path

    source\_URL: str

    local\_data\_file: Path

    unzip\_dir: Path

16. paste the configurationmanager() code in the config/configuration.py