

ML Project Assignment

Assignment Questions



Create and deploy a ML project by importing load_breast_cancer dataset from sklearn.load_dataset and apply the following:

1. Create a folder in which you want to create the project, after that use the git init and the necessary commands to create the specific Git repository.
2. Create a separate environment so that you do not mess up with your base environment.
3. Create the folder structure/directories and files using the python programme required for a ML project. You can refer the following project structure:

- src
 - __init__.py
 - logger.py
 - exception.py
 - utils.py
 - components
 - __init__.py
 - data_ingestion.py
 - data_transformation.py
 - model_trainer.py
 - pipeline
 - __init__.py
 - predict_pipeline.py
 - train_pipeline.py
- import_data.py
- setup.py
- notebook
- requirements.txt

After this update the created folders and files to your git repository by pushing from your end and add following files from github and pull it to your source code:

- README.md
- LICENSE
- .gitignore

4. Write the program for setup.py and the relevant dependencies in requirements.txt and generate egg.info folder.
5. Write the logging function in logger.py and exception function in exception.py file to be used for the project to track the progress when the ML project is run and to raise any exception when encountered.

7. In the notebook folder create a jupyter notebook inside it and do the following with the dataset:

- Exploratory Data Analysis
- Feature Engineering
- Model Training
- Selection of best model using metric

8. Write a separate python program in `import_data.py` file to load the mentioned dataset from `sklearn.load_dataset.load_breast_cancer` to your MongoDB.

9. In `data_ingestion.py` write a program to load the same dataset from the MongoDB to your system in `DataFrame` format.

10. Do the necessary feature engineering part in `data_transformation.py`.

11. Create the Machine Learning model in `model_trainer.py`.

12. Use Flask to deploy your project.

Note: Write any required utility function to be used while implementing the project code in `utils.py`.

Remember to use `.gitignore` to keep your environment, artifacts and necessary folders of the project to be safe. After each and every step do not forget to update your commits through `git` so that the project repository on `github` remains updated. Share the `github` link of this project repository and make sure to make it publicly accessible.