Modular Coding

1. Data Ingestion
2. Data Validation
3. Data Transformation
4. Model Trainer
5. Model Evaluation
6. Model pusher

**Data Ingestion**

**Workflow**

1. Us-visa 🡪 constants 🡪 copy the code

For all the components you need to assigned in the constant folder , so that instead of changing it in each file , we can change it from the single file

File contain important variables

1. In us-visa 🡪 configuration 🡪 create a mongo db connection.py file

All the folder and file path should be managed in the config

1. Us visa 🡪 create folder data\_access 🡪 create \_\_init\_\_.py and usvisa\_data.py 🡪
2. Us visa 🡪 entity🡪 config entity🡪
3. Us visa 🡪 entity 🡪 artifact (return type of the component . Here : data ingestion)🡪

Output from the components

Here train.py and test.py

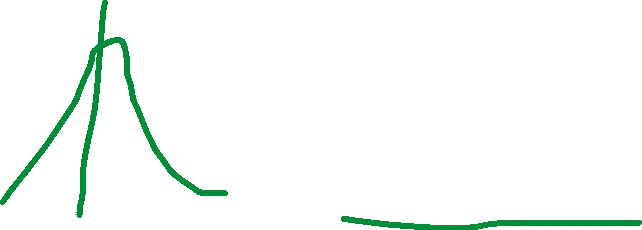
1. Us visa 🡪 Data ingestion.py 🡪
2. Training\_pipeline
3. Before that set the MONGODB\_URL in env variable or set MONGODB\_URL=”-------------"
4. End point

**Data Drift**

Data drift, also known as concept drift or dataset shift, refers to the gradual or abrupt change in the statistical properties of the data used to train a machine learning model. These statistical properties can include changes in the distribution of features, target labels, or the relationships between them.



No drift occurs



Drift occurs

Training and testing data should be follows the same distribution . To detect the data drift we can use the MLOPs tool Evidently AI