PRACTICAL 5

1. Setup Project Folder

```
mkdir prefect-practical
cd prefect-practical
```

2. Create and Activate Virtual Environment

```
python -m venv venv
.\venv\Scripts\activate
```

3. Install Libraries

```
pip install prefect scikit-learn pandas
```

4. Create Python file: `explain.py`

1. Setup Project Folder

```
mkdir prefect-practical
cd prefect-practical
```

2. Create and Activate Virtual Environment

```
 import time
import pandas as pd
from sklearn.datasets import make_classification
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score
from prefect import task, flow

@task
def extract_data():
    print("starting extract_data task")
    time.sleep(5)
    X, y = make_classification(n_samples=100, n_features=10, random_state=42)
    df = pd.DataFrame(X, columns=[f'feature_{i}' for i in range(10)])
    df['target'] = y
    df.to_csv('dummy_data.csv', index=False)
    print("task complete data saved")
    return 'dummy_data.csv'

@task
def transform_data(data_path: str):
    print("starting transform_data task")
    time.sleep(5)
    df = pd.read_csv(data_path)
    X = df.drop('target', axis=1)
    y = df['target']
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
    print("task done data splited")
    return X_train, X_test, y_train, y_test

@task
def load_model(X_train, y_train, X_test, y_test):
    print("starting load_model task")
    time.sleep(9)
    model = LogisticRegression()
    model.fit(X_train, y_train)
    preds = model.predict(X_test)
    acc = accuracy_score(y_test, preds)
    print(f"task done model trainned accuracy {acc}")
    return acc

@flow(name="Prefect Practical Exam Flow")
def run_pipeline():
    print("flow started hurry up")
    data_path = extract_data()
    X_train, X_test, y_train, y_test = transform_data(data_path)
    accuracy = load_model(X_train, y_train, X_test, y_test)
    print(f"flow cmplete accuracy {accuracy}")

if __name__ == "__main__":
    run_pipeline()
```

5. Run setup and training

```
python explain.py
```

6. In a new Terminal in same venv

```
prefect server star
```

Open your browser to `http://127.0.0.1:4200`

On Prefect Dashboard navigate to: `runs -> latest run`