Practical-2

Aim: Introduction to reproducible Machine Learning Operations

The aim of the practical is to get the hands-on experience of reproducing the machine learning operations at each stage. Students need to apply the following steps in the practical.

1. Ensure that the NumPy, scikit learn, and matplotlib libraries are available in your system. Create the requirements.txt file and make a note of the versions of these libraries.

```
C:\Users\Nakul>python -c "import numpy as np; print(np.__version__)"
1.23.5

C:\Users\Nakul> python -c "import sklearn; print(sklearn.__version__)"
1.3.2

C:\Users\Nakul>python -c "import matplotlib; print(matplotlib.__version__)"
3.7.1

C:\Users\Nakul>python -c "import pandas; print(pandas.__version__)"
1.5.3
```

2. Write a python code to import the Sample.txt data. Further, apply the following processes on the imported data.
    a. Scale the dataset.

```
import numpy as np
from sklearn.preprocessing import StandardScaler
import joblib
data = np.genfromtxt('sample.csv', delimiter=',')
scaler = StandardScaler()
scaled_data = scaler.fit_transform(data)
joblib.dump(scaler, 'scaler_object.joblib')
```

    b. Split the data:

```
from sklearn.model_selection import train_test_split
train_data, test_data = train_test_split(scaled_data, test_size=0.2,
random_state=42)
```

    c. Store the snapshot of the data as the NumPy file.

```
np.save('train_data.npy', train_data)
np.save('test_data.npy', test_data)
```

3. Apply the linear regression algorithm on the dataset and assess the prediction on the test dataset.
    a. Store the trained model into the local file system to ensure the reproducibility of the prediction. Import the model and the test dataset into another python file. Check whether the same prediction is obtained in the latter case.

```
from sklearn.linear_model import LinearRegression
import joblib
X_train = train_data[:, :-1]
y_train = train_data[:, -1]
model = LinearRegression()
model.fit(X_train, y_train)
X_test = test_data[:, :-1]
y_test = test_data[:, -1]
predictions = model.predict(X_test)
joblib.dump(model, 'linear_regression_model.joblib')
loaded_model = joblib.load('linear_regression_model.joblib')
loaded_predictions = loaded_model.predict(X_test)
if np.array_equal(predictions, loaded_predictions):
   print("Output : Predictions match!")
else:
   print("Output : Predictions differ!")
```

Output:

```
[Running] python -u "c:\Users\Nakul\Downloads\MLOPs\Practicals\app.py"
Output : Predictions match!
```