bytewise.

Bytewise Fellowship Program

DATA SCIENCE

Task 15

BWT- Data Science (Group1)

Submitted to: Mahrukh Khan

Submitted by: Usama Malik

Task: Advanced Numpy (reshape , concatenate, splitting , hsplit etc) - Appendx A

1. Reshape

Definition: Reshape allows you to change the shape of an array without changing its data.

Example:

```
import numpy as np

# Creating an array

arr = np.arange(12)

print(arr)


# Reshaping the array

reshaped_arr = arr.reshape(3, 4)

print(reshaped_arr)
```

2. Concatenate

Definition: Concatenate joins two or more arrays along an existing axis.

Example:

```
# Creating two arrays

arr1 = np.array([[1, 2], [3, 4]])

arr2 = np.array([[5, 6], [7, 8]])


# Concatenating along axis 0

concatenated_arr = np.concatenate((arr1, arr2), axis=0)

print(concatenated_arr)
```

```
# Concatenating along axis 1
concatenated_arr = np.concatenate((arr1, arr2), axis=1)
print(concatenated_arr)
```

3. Splitting

Definition: Splitting divides an array into multiple sub-arrays.

Example:

```
# Creating an array
arr = np.arange(16).reshape(4, 4)
print(arr)

# Splitting the array into two equal parts along axis 0
split_arr = np.array_split(arr, 2, axis=0)
print(split_arr)

# Splitting the array into two equal parts along axis 1
split_arr = np.array_split(arr, 2, axis=1)
print(split_arr)
```

4. HSplit

Definition: HSplit splits an array into multiple sub-arrays horizontally (column-wise).

Example:

```
# Creating an array
arr = np.arange(16).reshape(4, 4)
print(arr)
```

```
# Splitting the array into two equal parts horizontally
hsplit_arr = np.hsplit(arr, 2)
print(hsplit_arr)
```

5. VSplit

Definition: VSplit splits an array into multiple sub-arrays vertically (row-wise).

Example:

```
# Creating an array
arr = np.arange(16).reshape(4, 4)
print(arr)

# Splitting the array into two equal parts vertically
vsplit_arr = np.vsplit(arr, 2)
print(vsplit_arr)
```

6. DSplit

Definition: DSplit splits an array into multiple sub-arrays along the depth (depth-wise).

Example:

```
# Creating a 3D array
arr = np.arange(27).reshape(3, 3, 3)
print(arr)

# Splitting the array into three equal parts along depth
dsplit_arr = np.dsplit(arr, 3)
print(dsplit_arr)
```

7. Transpose

Definition: Transpose flips the dimensions of an array.

Example:

```
# Creating an array
arr = np.arange(12).reshape(3, 4)
print(arr)


# Transposing the array
transposed_arr = arr.T
print(transposed_arr)
```

8. Flatten

Definition: Flatten returns a copy of the array collapsed into one dimension.

Example:

```
# Creating an array
arr = np.arange(12).reshape(3, 4)
print(arr)


# Flattening the array
flattened_arr = arr.flatten()
print(flattened_arr)
```

9. Broadcast

Definition: Broadcasting describes how NumPy treats arrays with different shapes during arithmetic operations.

Example:

```
# Creating two arrays
```

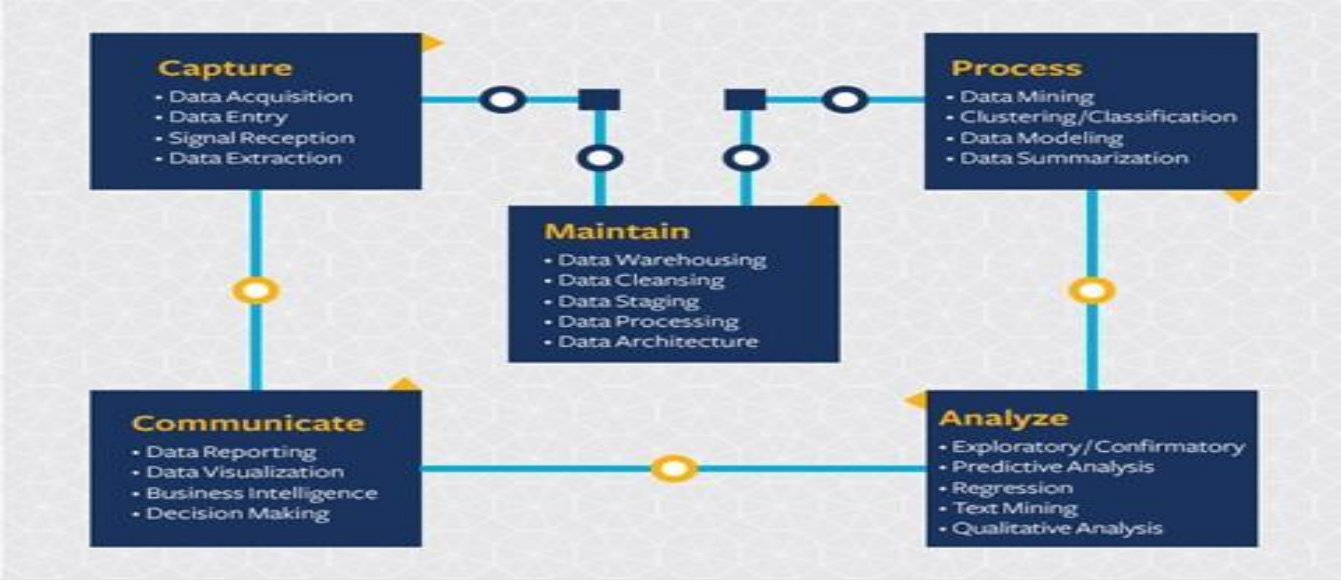
```
arr1 = np.array([1, 2, 3])

arr2 = np.array([[1], [2], [3]])


# Broadcasting addition

broadcasted_arr = arr1 + arr2

print(broadcasted_arr)
```

Data Science Life Cycle

At this point you've probably come to the realization that data science is a process. This process can be broken down into 5 stages:

- Capturing
- Processing
- Analysis
- Communication
- Maintenance

This lesson focuses on 3 parts of the life cycle: capturing, processing and maintenance.



□ Capturing Data:

The process of collecting raw data from various sources. For example, a survey collecting customer feedback.

□ Processing Data:

Converting raw data into a usable format. For example, cleaning and transforming survey responses into a structured dataset.

□ Maintaining Data:

Ensuring data remains accurate and up-to-date. For example, regularly updating a customer database with new information.

□ Storing Data:

Saving data in a structured way for future use. For example, storing customer records in a cloud database.

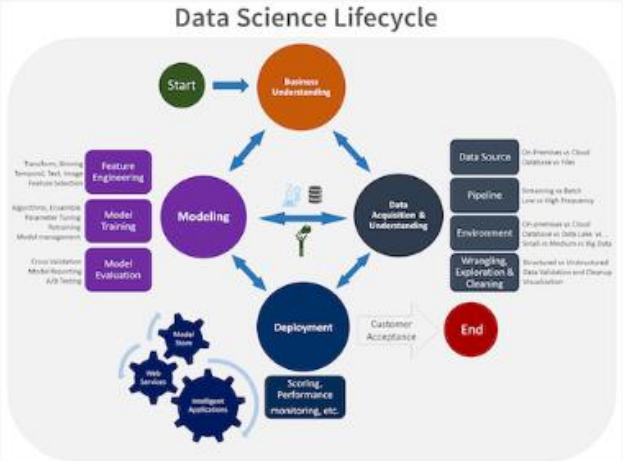
□ Managing Data:

Organizing and controlling data to ensure accessibility and reliability. For example, using database management systems to handle user data.

□ Securing the Data:

Protecting data from unauthorized access and breaches. For example, encrypting sensitive customer information stored in a database.

Team Data Science Process (TDSP)



Cross-industry standard process for data mining (CRISP-DM)