INFO5002: Intro to Python for Info Sys

Numpy

numpy.org

Slides created by: Zachary Doucet

LVX VERITAS VIRTVS
Northeastern University

Python lists are limiting

- Limited operations, heterogeneous data storage--operation that works at one index may not work at another.
- We will be using Numpy.

```
pip install numpy
```

```
import numpy as np
```

Array as primitive

- Numpy uses the array as a primitive—kind of like a Python list.

```
x = np.array([1,2,3,4,5,6])
```

- Arrays can be indexed and are mutable like python.

```
x[0] = 4
```

- Can also do slicing but be careful—returns a view instead of a copy. Mutating the view mutates the original!

```
y = x[:3]
y[0] = 6
```

Array Attributes

- arr.ndim: number of dimensions: 1 – vector, 2 – matrix, 3+ – tensor.
- arr.shape: returns the number of dimensions and data as tuple.
- arr.size: returns total number of elements.
- arr.dtype: returns the data type.

Functional array creation

- np.zeros(shape): create array of *shape* filled with 0's.
- np.ones(shape): create array of *shape* filled with 1's.
- np.empty(shape): create array of *shape* filled with random numbers (faster!).
- np.arange(start, stop, step): create vector of range from *start* to *stop* incrementing by *step*.
- np.linspace(start, stop, num=50): create vector of *num* evenly spaced numbers from *start* to *stop*.

Useful functions

- np.sort(arr): sorts in ascending order.
- np.concatenate((arr1, arr2, ...), axis=0): concatenates two or more arrays together along axis *axis*.
- np.reshape(arr, shape=shape): will reshape the array according to what you specify—must keep same # elems.
  - (12,) → (3,4) Good
  - (2, 4) → (3, 3) Bad

Conditional selection

- Can pass a conditional within brackets:

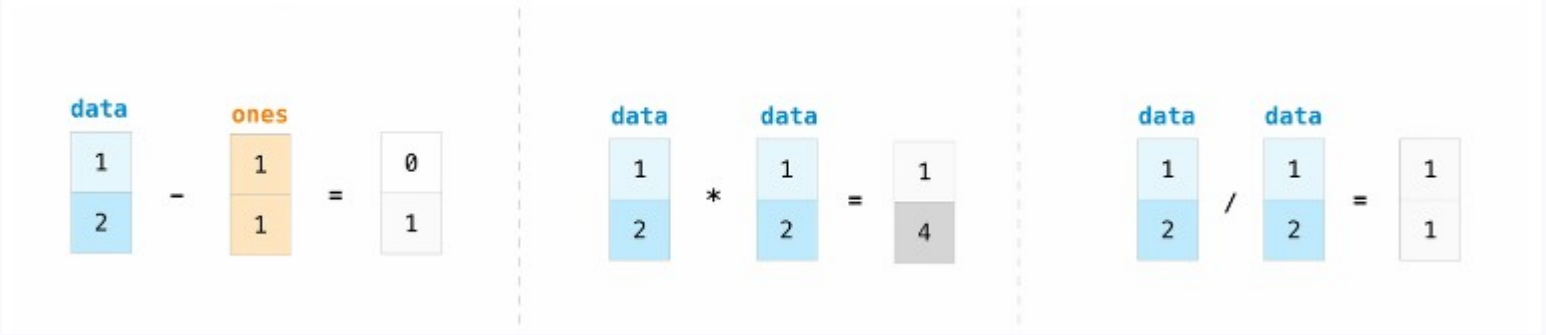
```
b = a[(a > 18) & (a < 25)]
```

- & is same as python's and and | is same as python's or.

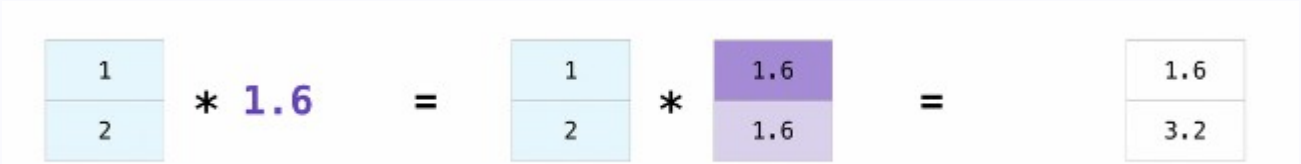
```
divisible_2_or_3 = a[(a%2==0)|(a%3==0)]
```

Operations

- +, -, /, * all do the operation element wise.



- If you try to perform an operation with two different shapes, it will attempt to broadcast to make it work.



Useful operator functions

- arr.sum(axis=-1): returns the sum of all the elements together along axis *axis*.
- arr.max(axis=-1): returns the largest along axis *axis*.
- arr.min(axis=-1): returns the smallest along axis *axis*.