Python: Scripting Language

Tassadaq Hussain

Riphah International University

Barcelona Supercomputing Center

UCERD Pvt Ltd

Outcome of this lecture

- Scripting Languages
- Installing python
- Importing and Using libraries
- Accessing Databases

Big Data Processing

1. Reading the data and cleaning it.
2. Exploring and understanding the input data.
3. Analyzing how best to present the data to the learning algorithm.
4. Choosing the right model and learning algorithm.
5. Measuring the performance correctly.

Computer Languages

- Programming Languages
  - Complier Dependent
  - Runs Directly on CPU
- Scripting Languages (HLL)
  - Java, Python, R, Scala
  - Interpreter Dependent
  - Runs with support of interpreter

Database support

- NumPy provides the support of highly optimized multidimensional arrays, which are the basic data structure of most state-of-the-art algorithms.
- SciPy uses those arrays to provide a set of fast numerical recipes.
- Matplotlib is probably the most convenient and feature-rich library to plot high-quality graphs using Python.

Learn by implementing

- import numpy, scypy, matplotlib
- tutorial
-
-
- sudo apt-get install python-numpy python-scipy python-matplotlib

Reading database

- a=np.genfromtxt('database.csv',delimiter=",")

runtime behaviors

- Import timeit
- timeit.default_timer()
-

Scipy

SciPy package	Functionality
`cluster`	Hierarchical clustering (`cluster.hierarchy`) Vector quantization / K-Means (`cluster.vq`)
`constants`	Physical and mathematical constants Conversion methods
`fftpack`	Discrete Fourier transform algorithms
`integrate`	Integration routines
`interpolate`	Interpolation (linear, cubic, and so on)
`io`	Data input and output
`linalg`	Linear algebra routines using the optimized BLAS and LAPACK libraries
`maxentropy`	Functions for fitting maximum entropy models
`ndimage`	n-dimensional image package
`odr`	Orthogonal distance regression
`optimize`	Optimization (finding minima and roots)
`signal`	Signal processing
`sparse`	Sparse matrices
`spatial`	Spatial data structures and algorithms
`special`	Special mathematical functions such as Bessel or Jacobian
`stats`	Statistics toolkit

- The toolboxes most interesting to deeplearning are
- scipy.stats
- scipy.interpolate
- scipy.cluster
- scipy.signal