

# The best Python tools for Machine Learning and Data Science

Python has many great libraries and frameworks that are good for coding and developing computer science. Quantrimang invites you to discuss some useful Python tools for both Machine Learning and Data Science applications.

Python programming language has many large libraries and frameworks that are convenient for coding and developing computer science. Python is a language that is famous for its unobtrusive simplicity, easy-to-learn code, easy to read, logical syntax and concise, while **Machine Learning** involves extremely complex algorithms and multi-stage workflows. So here, Python's brief and easy logic plays an important role in saving time for developers.

On the other hand, when it comes to **Data Science**, Python also has special packages for these field tasks like SciPy, NumPy or Pandas that facilitate data analysis and can be easily Integrated with web applications.



In addition, Python is actually an open source language, you can freely use and distribute Python, even for commercial use. Thanks to that, Python has a lot of high quality resources and materials, and a community of active developers willing to provide advice and support at all stages of the development process.

So Quantrimang invites you to discuss some useful Python tools for both Machine Learning and Data Science applications.

## Python tool for Data Science



## 1. NUMBA

**Numba** is an NumPy-aware open source compiler that compiles Python syntax into machine code using the LLVM compiler funded by the Anaconda. Numba applications in Data Science increase the speed of compiling code with NumPy Array. Provided some Annotation, Python code can be optimized to achieve the same performance as C, C ++ and Fortran without having to change the language or interpreter.

## 2. CYTHON

**Cython** is a variant of C of Python. It can be said that it is a parent of Python, capable of creating standard Python modules, greatly improving execution speed and performance. Basically, it is designed as an extension of C for Python to compile Python code into C / C ++ code and is used in Jupyter notebooks through inline annotations.

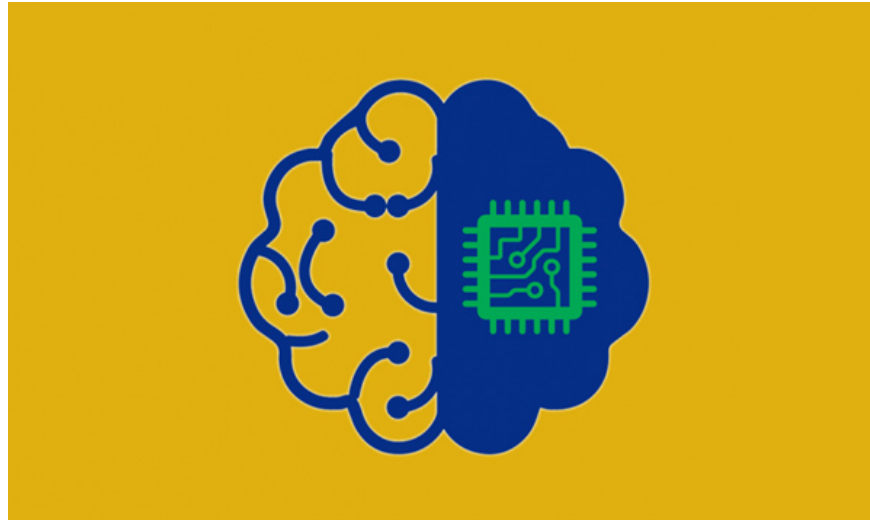
## 3. DASK

**Dask** is a flexible library for parallel computing in Python. When using Numpy or Pandas, you sometimes face the problem of processing data in RAM, where Dask is easy to handle because it expands the interfaces to larger or memory-dispersed environments. Can run on local or minimized computers to run on a cluster.

## 4. SCIPY

**SciPy** is an open source library of algorithms and mathematical tools for Python, built on NumPy array objects that constitute NumPy stacks including tools like Pandas, SymPy and Matplotlib. SciPy provides quite a lot of modules that calculate from linear, integral, differential, interpolation to image processing, fourier transform .

## Python tool for Machine Learning



## 1. SCIKIT-LEARN

**Scikit-learn** (sklearn for short) is an open source library for Machine Learning and is also used in Data Science. This is a very powerful and popular tool for the Python community, designed on NumPy and SciPy platforms. Scikit-learn contains most of the most modern Machine Learning algorithms, along with documentations, that are always up to date. This tool provides easy API usage and random search. But the main advantage in using Scikit-Learn, is speed while making different assessments in the dataset.

## 2. KERAS

**Keras** is an open source library written in python for neural network. Keras is a high-level API, developed to make the deep learning models as fast and easy as possible for research, with the MIT license for open source software. This tool can be used in conjunction with the famous Deep Learning libraries such as TensorFlow, CNTK, Theano.

Keras has several advantages such as:

1. Easy to use, build modules fast.
2. Can run on both CPU and GPU
3. Support to build CNN, RNN and can combine both.
4. Easy scalability and work with Python.

## 3. THEANO

**Theano** is an open-source Python library that supports arithmetic operations that can run on CPUs or GPUs, used to build and develop Deep Learning models. Theano provides very convenient structure and model adjustment methods to use on Numpy library functions to calculate, can run on GPU architecture outside CPU to be effective. Theano also created C code flexibly, extensive unit testing and self-verification, optimizing speed and stability. This is the first library to build and develop an artificial neural network learning model using deep learning techniques since 2007 and is considered as a technology standard for Deep Learning technology in the research and development community.

This is the list of Quantrimang. If you think there is an important tool that has been missed on this list, please comment below to add it to Quantrimang.

You finished reading the article "**The best Python tools for Machine Learning and Data Science**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.

---