

Python ライブラリ事典 101

久保 幸夫

現在Python言語は、データの統計処理やデータ解析、さらには機械学習などへの活用が進んでいるため、ビッグ・データや人工知能のブームとも相まって、もっとも注目されている言語と言われています。

Pythonのライブラリには多くのものがあります(表1)。その一部を表2～表10に示します。

本記事は2016年10月号特集 第2章「IoT/データ解析にピッタリ! おすすめPythonライブラリ事典101」を、筆者: 齊藤 直希氏の許可を得て、加筆修正したものです。

表1 Python向けに提供されているライブラリは多数ある

| 表番号 | 項目 | 詳細 | 紹介項目数 |
|-----|------------------|-----------------------------------|-------|
| 2 | 科学/ データ解析 | 数値計算、信号処理、統計処理など | 17 |
| 3 | 機械学習・ 認識 | 機械学習、深層学習、ニューラル・ネットワークなど | 13 |
| 4 | マルチメディア | 音声、画像、動画など | 8 |
| 5 | ネットワーク、 通信 | 各種プロトコル実装(SSH、SMB)、パケット解析、SNSなど | 23 |
| 6 | コンピュータ・ ボード向け | GPIO、シリアル、ラズベリー・パイ、ロボットなど | 10 |
| 7 | Web開発 | Webフレームワーク、HTMLパーサなど | 9 |
| 8 | 画面制御 | GTK+などのGUI、OpenGLなどのグラフィックス・ライブラリ | 6 |
| 9 | データベース | PostgreSQL、MySQLなどのデータベース・ドライバ | 8 |
| 10 | 自然言語処理 | 形態素解析などの自然言語処理 | 7 |
| 11 | ライセンス | ライブラリでよく使われるライセンス | 13 |

①科学 / データ解析

データの統計や機械学習には、数値計算が欠かせません。Pythonでは数多くの科学計算や数値演算などのライブラリが提供されており、科学技術計算や統計解析の用途に加え、人工知能関連でよく利用されています(表2)。

特にNumPyは、古くからあるベクトルや行列演算を行う数値演算ライブラリで、TensorFlowやChainer(表3)などの機械学習のフレームワークの中でもよく使われています。

また、NumPyから派生して、科学技術計算を追加するSciPyや、グラフ描画機能を追加するmatplotlib、記号計算を強化するSymPyなどのライブラリがあります。

TheanoやCuPyはGPUを使用した高速な数値計算

を目指したライブラリで、NumPyの代替として使うことができます。TensorFlowは、データフロー・グラフを用いた数値計算フレームワークですが、ニューラル・ネットワークなどの記述にも適しているため、機械学習のフレームワークとしても使われています。

データ解析を便利にするライブラリも豊富で、pandasなどのデータ解析を支援するライブラリがあります。また、他の言語をPythonから利用するためのライブラリもあり、例えば、Oct2Pyは統計処理やデータ解析でよく使われているR言語を利用するためのライブラリです。