

機械学習のための Python入門

～Python・PyDataについて～

TFUG Utsunomiya #12

2018/06/17 Akimasa KIMURA



PythonとPyData

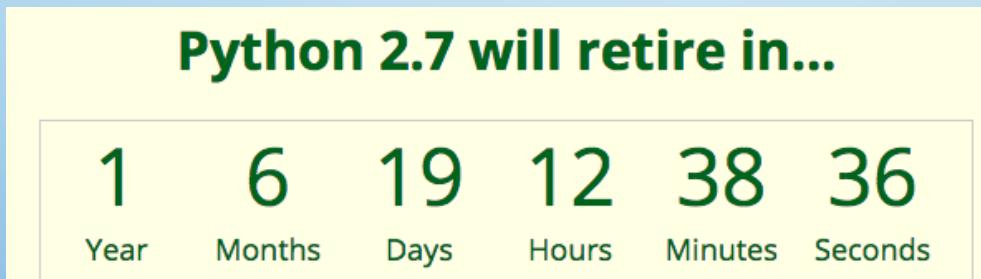
- Python
- Pythonのエコシステム
 - サードパーティ
 - コミュニティ
 - 広がる利用用途
- PyDataとは何か

Pythonとは？

- 1991年にGuido van Rossum[1]によって開発された
- 2018年現在 Pythonのバージョン：
3.x系は**3.6.5**
2.7.x系は**2.7.15**
- 2.7.x系のサポートはあと1年ちょっと[2]



[1]



[1]

<https://ja.wikipedia.org/wiki/%E3%82%B0%E3%82%A4%E3%83%89%E3%83%BB%E3%83%B4%E3%82%A1%E3%83%B3%E3%83%AD%E3%83%83%E3%82%B5%E3%83%A0>

[2] <https://pythonclock.org/>

Pythonとは？

- Python自体は言語仕様で
 - <https://docs.python.org/3/reference/grammar.html>
- 実装はたくさんある
- 普段よく使うのは、リファレンス実装である**CPython**
- JVM上で実装された：**Jython**
- Pythonで実装されたPython：**PyPy** etc

Pythonとは？

- マルチプラットフォーム対応：
 - Windows
 - OSX / macOS
 - iOS
 - Android
 - Unix / Linux etc
- 詳しくはPython Japanの環境構築ガイドを参考
 - <https://www.python.jp/install/install.html>

Pythonのエコシステム

- **PyPi** : Pythonのライブラリ・モジュールのアーカイブ
 - 公開されたPYTHONのモジュールをインストールして利用できる
 - 例: `pip install tensorflow`
- **PEP** : Python Enhancement Proposal
 - 提案やアイディアはオープンな形で議論される
- **PyCon** : カンファレンス
 - 世界中で開催
 - EuroPython, Pycon Asia Pacific etc



データサイエンスで使われるPython

- なぜPythonなのか？
- <https://www.quora.com/Why-is-Python-a-language-of-choice-for-data-scientists>

データサイエンスで使われるPython

- よく使われるライブラリ群：
- **numpy** : ベクトル・行列演算用パッケージ
- **scipy** : 科学技術計算用パッケージ
- **matplotlib** : グラフ描画用パッケージ
- **scikit-learn** : 機械学習用パッケージ
- **pandas** : データ解析用パッケージ

Pythonでデータサイエンスを はじめてみよう

- データサイエンスのためにPythonを使う方法
- **Anaconda**
- **Enthought Canopy**
 - すでにnumpyやpandas、matplotlib etcが組み込まれている
- **Colaboratory**
 - Googleが公開しているJupyter Notebookの環境

機械学習のための Python入門

～Python・PyDataについて～

TFUG Utsunomiya #12

2018/06/17 Akimasa KIMURA



文法の仕様

```
# Start symbols for the grammar:
#   single_input is a single interactive statement;
#   file_input is a module or sequence of commands read from an input file;
#   eval_input is the input for the eval() functions.
# NB: compound_stmt in single_input is followed by extra NEWLINE!
single_input: NEWLINE | simple_stmt | compound_stmt NEWLINE
file_input: (NEWLINE | stmt)* ENDMARKER
eval_input: testlist NEWLINE* ENDMARKER

decorator: '@' dotted_name [ '(' [arglist] ')' ] NEWLINE
decorators: decorator+
decorated: decorators (classdef | funcdef | async_funcdef)

async_funcdef: ASYNC funcdef
funcdef: 'def' NAME parameters ['->' test] ':' suite

parameters: '(' [typedargslist] ')'
typedargslist: (tfpdef ['=' test] (',' tfpdef ['=' test]))* [',' [
    '*' [tfpdef] (',' tfpdef ['=' test])* [',' ['**' tfpdef [',']]]
    | '**' tfpdef [',']]]
    | '*' [tfpdef] (',' tfpdef ['=' test])* [',' ['**' tfpdef [',']]]
    | '**' tfpdef [','])
tfpdef: NAME [':' test]
varargslist: (vfpdef ['=' test] (',' vfpdef ['=' test]))* [',' [
    '*' [vfpdef] (',' vfpdef ['=' test])* [',' ['**' vfpdef [',']]]
    | '**' vfpdef [',']]]
    | '*' [vfpdef] (',' vfpdef ['=' test])* [',' ['**' vfpdef [',']]]
    | '**' vfpdef [',']]
)
vfpdef: NAME
```