

Minicondaと仮想環境

INIAD

データサイエンスのための環境構築

- pythonには、データサイエンス用のライブラリ(モジュール)が多数用意されており、importすることで手軽に使うことができます。
- 機械学習法のライブラリは、更新ペースが速い、頻繁に仕様変更がある、ライブラリ間のバージョンの組み合わせによっては動作しないことがある、などのために、CS2と同様に仮想環境を構築し、その上で作業することが多いです。



(復習) 仮想環境とは

- 仮想環境は1台のPC上に、pythonとライブラリの複数の組み合わせを 共存させることができる仕組みです。
- CS2では「venv」を利用して仮想環境を構築しましたが、CS3では、 より複雑なライブラリ間のバージョンコントロールを実現し、データ サイエンス用に広く用いられている「Miniconda」を利用します。

システム

python ライブラリ …

Miniconda

(base) python

仮想環境1

python ライブラリ

ライブラリ

仮想環境2

python

ライブラリ

ライブラリ・

システムのpython やライブラリに影響を与えず、独立 したpython+ライ ブラリ環境を複数 構築できる

INIAD

Miniconda上の仮想環境

- Minicondaのインストールと仮想環境の構築の手順は、MOOCsのCS3-00回「データサイエンスのための仮想環境構築」をみてください。
- 今回構築した仮想環境「ds2025」のpythonやJupyterLabなどは、CS2で使ってきた venvを用いた仮想環境とは別のものです。

● 昨年度の「ds2024」を作ってある人も、必ず今年度用の「ds2025」をあらためて作成し

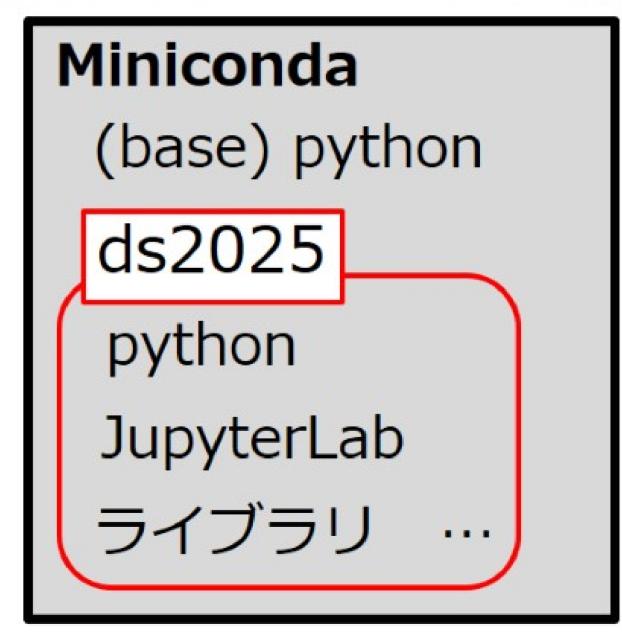
てください (配布スクリプトが動かないことがあります)。

システム

python ライブラリ … cs2024_algo/venv

python JupyterLab

ライブラリ …





ds2025仮想環境にインストールされている主なライブラリ

| ライブラリ | 説明 |
|----------------------------|------------------------|
| numpy (ナムパイ) | 多次元配列などの数値演算 |
| scipy (サイパイ) | 科学技術計算 |
| pandas (パンダス) | 表形式のデータの解析 |
| matplotlib (マットプロットリブ) | グラフ描画 |
| seaborn (シーボーン) | グラフ描画 |
| statsmodels (スタッツモデルズ) | 回帰などの統計解析 |
| scikit-learn (サイキット・ラーン) | 機械学習 |
| tensorflow (テンサー(テンソル)フロー) | 深層学習 |
| h5py (エイチファイブパイ) | HDF5形式バイナリデータ処理(深層学習用) |
| pillow (ピロー) | 画像処理 |

jupyterlab(ジュパイターラボ, ジュピターラボ): 実行環境「JupyterLab」 ※この他にも付随して多数のライブラリが自動的にインストールされています