

はじめにすること

- [ここ](#)からcsvファイルをダウンロード
- ファイル名を”muki.csv”に変更

はじめにすること

- [ここ](#)からsupporting info (pdf)ファイルをダウンロード
- Table S1,2をもとに、下記のような表を作成
- ファイル名を” organic synthesis.csv”として保存
- (著作権の関係で、ファイル直接配布はできません)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
Solvent	MW	tPSA	MP	BP	Density	Vdensity	Vpressure	LogP	Viscosity	Hvapor	Tension	IP	Epsilon	HOMO	LUMO	DeltaH-L	Dipole	Yield
Et2O	74.1	9.2	-116	35	0.71	2.6	538	0.89	0.24	27.85	17.1	9.53	4.3	-0.371	0.192	0.563	1.08	56
acetone	58.1	17.1	-94	56	0.79	2	231	-0.24	0.32	30.99	23.7	9.69	20.7	-0.369	0.103	0.472	2.94	53
M.OM	44.1	22.2	44	22	0.72	1.42	22.2	0.24	0.25	22.22	22	12.2	27.5	0.452	0.151	0.227	2.22	42

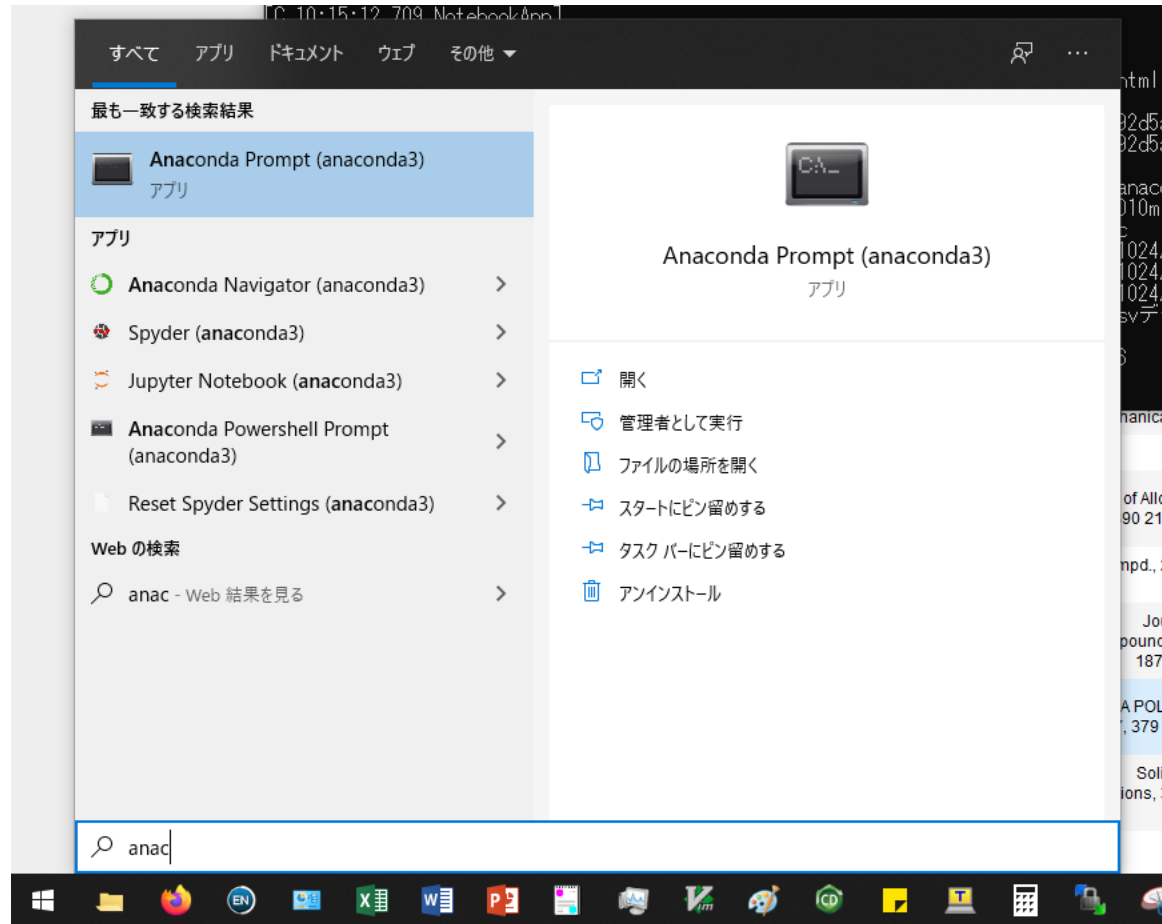
TODO

- サンプルファイルの実行
 - 因果探索
 - Random forest
 - Lasso
 - Gaussian process
 - 次の実験条件の計算
- 自分のデータで試す (残りの時間)
 - 上記のサンプルコードを自分のデータで実行してみる
 - 色々と自分で試してみる (他の回帰モデルなど)
 - (細かいテクニックなどはマテリアルズ・インフォマティクスβで扱います)

位置づけ

- 教師あり学習
 - 深層
 - 各種ニューラルネットワーク(畳み込み、回帰、グラフ...)
 - 最近の技術 (深層強化学習, BERT, ...)
 - 通常
 - 線形 (多重共線性の回避)
 - Lasso
 - Ridge, Elastic net, Bayesian ridge, ...
 - 非線形
 - Random forest, Gradient boosting, ...
 - Gaussian process, Support vector machine
- 教師無し学習
 - 因果探索
 - バイズ推論、敵対生成、...
 - クラスタリング (PCA, tSNE, UMAP,...)
 - 深層学習 (オートエンコーダ, 敵対生成ネット, ...)
- 精度検証
 - 訓練・検証データセットの分割 (扱っていない内容)

参考: ライブラリインストール



ライブラリのインストール

```
conda install -c ankurankan pgmpy -y
```

```
conda install -c anaconda networkx -y
```

と入力してエンター

(コピペがおすすめ。Proxy環境は不可)

因果探索に必要です。