**材料の物性予測　提出コードの実行手順説明**

**基本方針**

Geo-CGNNの論文の著者が公開しているコードを実行するだけ。

https://github.com/Tinystormjojo/geo-CGNN

ただし、nishikaの公開しているデータセットのフォーマットではgeoCGNNは直接読み込んで実行することができないため、①と④でその点を微調整した。

**手順**

①**geo CGNNのためのデータの成形**

このフォルダ内の

cifファイル、ターゲット変数のcsvの作成.ipynb

を用いて、nishikaのデータセットからローカル環境下で個々の結晶構造の情報が記されたものである約１０万個のcifファイルと物資のidと目的変数の組が記録されている表であるtargets\_MP\_Ef.csvを１つ作成する(ローカル環境でcifファイルを作成する理由は、google colabのdrive上に大量のファイルを保存すると重くなるため)。これによりgeo CGNNが読み込めるデータ形式になる。

**②結晶グラフの作成**

つぎに、cifファイルを著者のページからAtomGraph.py に読み込ませ実行する。これは、結晶のグラフを作成するプログラムで、学習設定に応じて引数を指定して実行する。

例python AtomGraph.py --data\_dir ./database/cif --name\_database MP\_test1 --cutoff 8 --max\_num\_nbr 12 --compress\_ratio 1

**③グラフニューラルネットワークの学習と推論**

②で得られたグラフと①で作成したtargets\_MP\_Ef.csv を使い、著者のgithubページにアップロードされているprocess\_geo\_CGNN.py を、google colab proの環境下で実行する。

ここでも学習設定に応じて引数を設定して実行する。

ここではネットワークの学習と推論を行う。

**④nishikaに提出するためのデータの後処理**

③で得られた推論データである、test\_predictions\_6\_8\_12.csv

を、このフォルダ内にある提出用.ipynbで読み込みセルを実行し、nishika指定のフォーマットに直し提出する。

**学習設定**

本コンペで提出した分析での学習設定を示す。

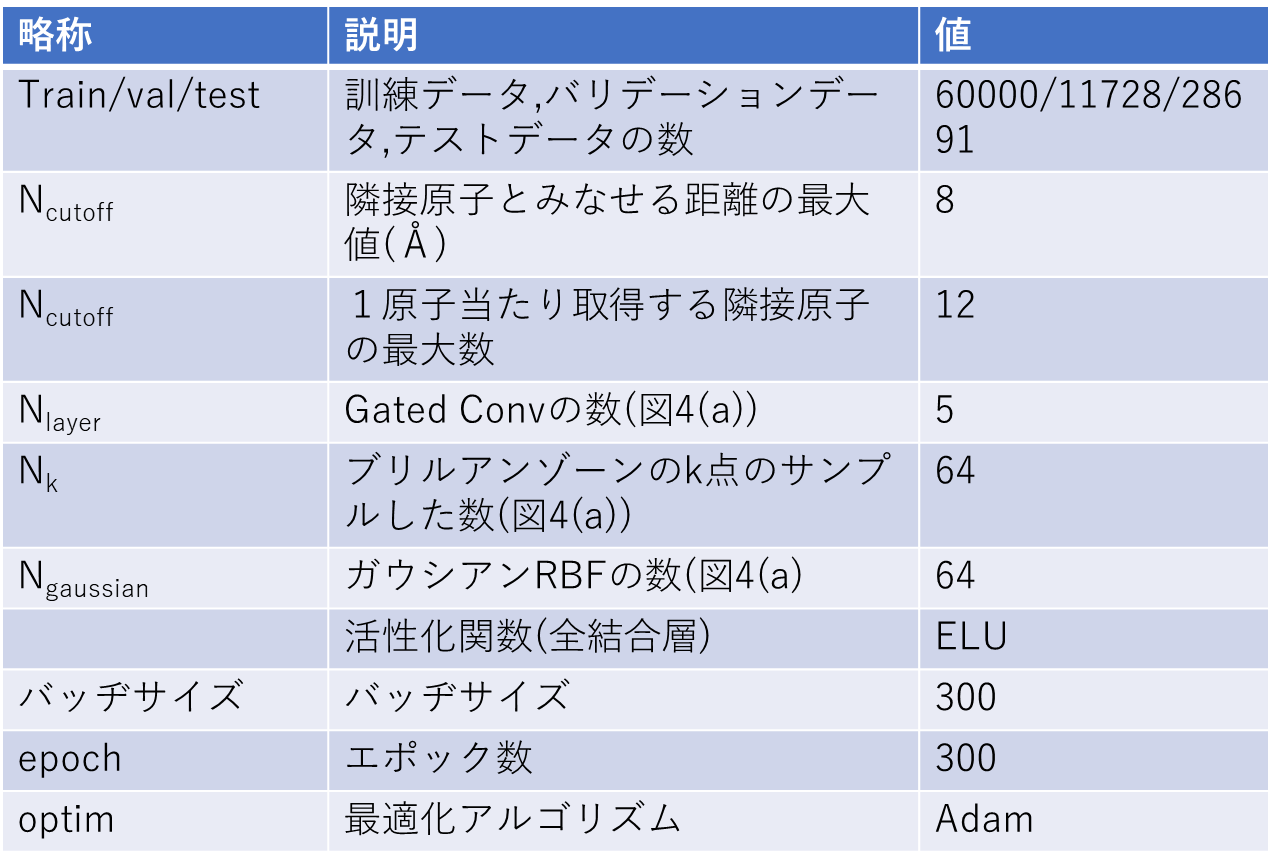


表１　学習設定