

## Colaboratory の教育・研究に対する有効性

近藤正紀<sup>1)</sup>

1) 新潟医療福祉大学 医療情報管理学科

【背景・目的】 人工知能開発やデータサイエンスが流行している現在、そのプログラミング言語として Python を使用することが一般的になっている。

Python は習得しやすい言語であるとされるが開発環境を整えることが難しいとされてきた。近年は環境構築ツール Anaconda が整備され容易になってはきたが、自身が所有するハードウェアに GPU を搭載して開発環境を構築すると、コストも労力も大きいのが現実である。

一方で、Google 社の提供する Colaboratory という IaaS では Python 環境を無料で使用することができる。

本稿は、Colaboratory を教育および研究で使用した場合の有効性について検討する。

【方法】 現在までに四つの講義と三つの研究プロジェクトで Anaconda と Colaboratory を使用した経験に基づき、Colaboratory とスタンドアロン PC で同様の環境を構築および維持する場合の困難度を比較検討する。本稿ではスタンドアロンを「ネットワークに接続していない」ではなく、「ネットワークに接続しなくても動作可能」という意味で用いる。

なお、本研究は倫理審査の対象外のものであり、関連する利益相反はない。

【結果】 Colaboratory は Jupyter Notebook をユーザーサイドの Web ブラウザーから操作するため、コード欄で各種の UNIX コマンド実行することによってハードウェア/ソフトウェアの構成を得ることができる。

表 1 Colaboratory の主な構成 (2019 年 9 月 1 日時点)

OS	Ubuntu 18.04.3 LTS
CPU	Intel Xeon 2.20GHz 2 cores
GPU	NVIDIA Tesla K80
主記憶	12.72 Gi バイト
ディスク	327.69 Gi バイト

(主記憶、ディスクは GPU 使用時)

Colaboratory はこの環境を Google ドライブでのわずかな操作で構築できる。スタンドアロン PC においてプログラムの開発環境を構築するには、インストールパッケージの入手とその展開を経た上で、カスタマイズを行うのが一般的である。

【考察】 スタンドアロン PC における環境構築は、ある程度の経験が無いと失敗することが多い。講義中に学生が

環境構築を行うと約半数の学生は 1 回以上構築をやり直すということを見てきた。また、構築に要する時間は長い場合 1 時間以上必要であるため、初回に失敗すると講義時間内に環境構築が完了しないというケースも多い。特に、GPU を使用する環境を構築する難度は高い。

これに対し Colaboratory は既に Jupyter Notebook が動作する状態で提供され GPU も利用できるため、Python を用いる場合は何もしなくても良い。稀に追加パッケージが必要となるが、追加に失敗しても Colaboratory の仮想コンピューター環境は容易に環境を抹消できることから、再構築は容易である。スタンドアロン PC では環境構築に失敗すると再構築前に不要データ群の削除が必要であり、削除漏れ等が起こると再び失敗する可能性が高くなる。環境構築が容易である点は利用者が本来の仕事に専念できることにつながり、教育・研究に対する有効性は高いと考える。

Python 以外の言語、例えば C を使用する場合は、ソースコードをユーザーサイドで作成し、それを Colaboratory に転送してコンパイルするというクロス開発を行う必要があるが、Ubuntu (UNIX) の CLI 環境での開発経験があれば困難ではないと考える。

アプリケーションで処理するデータは、IaaS による環境ではクラウド側のファイルシステムに転送しなければならない。Colaboratory は Google ドライブをファイルシステムにマウントできることから、ユーザーサイドのデータをファイル転送を意識せず扱うことができる。ただし、2019 年 5 月 31 日に『Team Drives』のパス名が『Shared drives』に変更されたため、前日まで動作していたプログラムが動作しなくなるというトラブルが発生した。このようなトラブルに対処するためには OS 等についてある程度のスキルが必要であり、さらに今後もこのような仕様変更に注視していく必要があると考える。

欠点として、インターネットへの接続が必須であること、90 分間操作が無いと接続が切れること、仮想コンピューターの 1 回の保持時間が 12 時間であることが挙げられる。接続が切れると最悪の場合演算結果を失うことになるため、Web ブラウザー側での細工や中間結果を保存し、それを読み込んで演算を継続するプログラミングテクニックが必要である。これらの対処は若干難度が高いが、一度設定してしまえばコードの再利用は容易であるため解決可能であり、利用者が本来業務に専念可能という前述の有効性を損なうものではないと考える。

【結論】 Google Colaboratory は高価なハードウェアを所有することなく、また環境構築の手間をかける必要なく、Python による教育研究環境を整えることができる。そのため本来業務に専念できる点で、非常に有効なツールであるといえる。