

1. 担当 PM

竹迫 良範（株式会社リクルート データプロダクトユニット ユニット長）

2. クリエータ氏名

妹尾 卓磨（慶應義塾大学）

3. 委託金支払額

2,460,975 円

4. テーマ名

非専門家でも手軽に使えるデータ駆動型深層強化学習ライブラリの開発

5. 関連 Web サイト

- d3rlpy : <https://github.com/takuseno/d3rlpy>
- d4rl-pybullet : <https://github.com/takuseno/d4rl-pybullet>
- d4rl-atari : <https://github.com/takuseno/d4rl-atari>
- MINERVA : <https://github.com/takuseno/minerva>

6. テーマ概要

本プロジェクトでは、データ駆動型深層強化学習アルゴリズムを誰でも簡単に利用できるようにするライブラリ d3rlpy を開発する。

深層強化学習はアカデミアや産業界での注目度の高さに対して、実際のプロダクトや応用事例が非常に少なく、原因として実問題への応用の難しさと実装難易度の高さが考えられる。また、深層強化学習アルゴリズムの実装が困難である要因として、非研究者向けに設計されたライブラリが存在しないことが挙げられる。深層強化学習をスクラッチで実装する場合、実装者には深層学習と強化学習の知識に加えて、論文の数式を解釈して実装に落とし込む能力が要求される。また、すでに多くの深層強化学習ライブラリが公開されているが、全てが機械学習の専門家・研究者向けに設計されているため、機械学習に精通していない利用者が使いこなすのが非常に難しいといった背景がある。

7. 採択理由

データ駆動型の深層強化学習ライブラリをオープンソースソフトウェアとして開発するという野心的な提案である。シミュレータ上でのオンライン学習を前提とした既存の強化学習ライブラリはドメイン特化のインテグレーション作業を必要とし、汎用性が低く、学習にも膨大な時間がかかる。本提案は少ないデータで強化学習ができる汎用のフレームワークを開発することに大きなモチベーションがある。研究分野の最先端のアルゴリズムを誰でも手軽に使えるような状態にできれば、実社会への AI 実装がさらに加速すると期待され採択した。

8. 開発目標

本プロジェクトでは、強化学習の応用の難しさと実装難易度の高さの両方を解決するようなライブラリ・ソフトウェアを開発することで、より多くの人が強化学習を使いこなせる世界を実現することを目的とした。

具体的には、実問題への応用の難しさに対して、近年注目が高まっているデータ駆動型強化学習のサポートを行うこととした。データ駆動型強化学習とは事前に集められたデータセットのみを用いてオフラインで学習を行うアプローチで、今まで強化学習を適用できなかった問題を扱うことができる。また、実装難易度の高さに対しては、ユーザが強化学習の非専門家であっても手軽に使える API を提供することで解決を行うこととした。

9. 進捗概要

(1) d3rlpy : 深層強化学習ライブラリ

d3rlpy は Python 向けの深層強化学習ライブラリである。最大の特徴としてデータセットのみからオフラインで学習ができるデータ駆動型強化学習アルゴリズムを多数実装している。非常に短いコードで最先端のアルゴリズムが利用できる。また、従来のオンライン学習もサポートしているため、あらゆる強化学習問題に利用することができる。

さらに、d3rlpy では強化学習アルゴリズムの原論文で示されている性能を超えるための機能が多数実装されており、本ライブラリ単独で専門家が実装したアルゴリズムよりも遥かに高い性能を得ることができる。他にも実際の製品に利用されることを想定した ONNX エクスポート機能や、データセットを Cython によって高速に処理するなど、利用者の利便性を最大化するための機能が盛り込まれている。

(2) d4rl-pybullet / d4rl-atari : データ駆動型強化学習向けデータセット

本プロジェクトでは d3rlpy の評価と強化学習コミュニティへの貢献を目的として、データ駆動型強化学習向けのデータセット d4rl-pybullet と d4rl-atari の 2 種類を公開した。

(3) MINERVA : 深層強化学習 GUI ソフトウェア

さらに、本プロジェクトではコードを全く書かずに最先端のデータ駆動型深層強化学習の利用を可能にした GUI ソフトウェア MINERVA を開発した(図 1)。

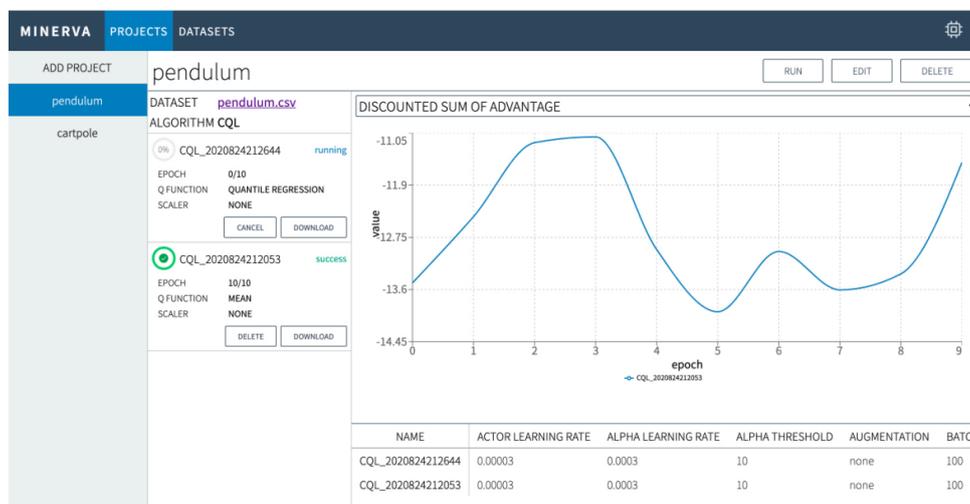


図 1 データ駆動型深層強化学習 GUI ソフトウェア MINERVA

MINERVA は Web ベースのソフトウェアであり、フロントエンドは React によって構築し、バックエンドの強化学習アルゴリズムに d3rlpy を利用している。

10. プロジェクト評価

深層強化学習は非常に活発な研究分野であるため、現在までに多くのライブラリが公開されてきたが、データ駆動型アルゴリズムをサポートしたライブラリは d3rlpy が世界初である。さらに、従来のオンラインアルゴリズムもサポートしているため、様々な強化学習問題に対して利用することができる。従来の深層学習ライブラリは研究者向けに設計されているため、拡張性を追求している反面、ユーザに多くのコードを書かせることが多かったが、d3rlpy では拡張性を犠牲にせず、最小限のコードで利用できるように設計されている。

GUI のみで機械学習を行うツールやサービスは非常にたくさん提案されてきたが、GUI のみで深層強化学習を行うツールは MINERVA が初めてである。

11. 今後の課題

本プロジェクトの開発物は GitHub 上で公開しながら開発を行っており、d3rlpy は 130 を超えるスター、38 件の issue 報告、13 件のプルリクエストが届いており、実際にこれらのライブラリを必要としているユーザ獲得の成果が得られている。ただ、MINERVA を使うシーンを考えるとまだまだ現場にいる専門家以外へ情報が行き届いていないこともあるため、非専門家へのアプローチが今後の課題である。