

Павло Мриглоцький. Рев'ю “ЕВОЛЮЦІЯ АГЕНТІВ НАВЧАННЯ З ПІДКРІПЛЕННЯМ ЗА ДОПОМОГОЮ ГЕНЕТИЧНОГО АЛГОРИТМУ”

Вступ

Стаття Артема Волокити та Богдана Гереги «ЕВОЛЮЦІЯ АГЕНТІВ НАВЧАННЯ З ПІДКРІПЛЕННЯМ ЗА ДОПОМОГОЮ ГЕНЕТИЧНОГО АЛГОРИТМУ» розглядає використання генетичного алгоритму машинного навчання з підкріпленням. Автори статті показали, що генетичний алгоритм працює більш «природньо» та точність навчання буде вищою.

Основна мета дослідження показати процес побудови нейронної мережі навчання з підкріпленням за генетичним алгоритмом.

Методологія

Теоретичне обґрунтування

На початку статті автори наводять теоретичний кістяк освітлення поставленої теми, даючи означення як поняттю «навчання з підкріпленням», так і «генетичного алгоритму», показуючи засадничу різницю з алгоритмом «градієнтного спуску». Також перераховуються кроки генетичного алгоритму та даються означення основним компонентам.

Аналіз

Далі автори ділять поставлену проблему на частини:

- Налаштування генетичного алгоритму
- Побудова нейронної мережі
- Опис очікуваного результату

Кожна частина оснащена означеннями основних понять, принципів роботи та призначення кожного кроку.

Синтез

Розглянувши деталі роботи моделі навчання з підкріпленням за допомогою генетичного алгоритму, автори описали будову нейромережі, що імплементує ці алгоритми. Звісно, що автори ілюструють опис роботи нейромережі рисунками, схемами та поясненням параметрів.

Далі автори наводять виміри ефективності роботи генетичних агентів після кожної ітерації. Вимірювання супроводжуються візуалізацією за допомогою графіків.

Методологічні недоліки

До методологічних недоліків я можу сміливо віднести відсутність жодних прикладів коду як самих нейронів, так і кроків алгоритму. Нема нічого кращого для розуміння роботи алгоритмів, аніж вивчення на прикладах шматків коду.

Також наводячи виміри ефективності, автори не описують методології вимірювання, що не дає оцінити коректність самої процедури.

Результати

З даних, наведених у статті бачимо, що генетичний алгоритм навчання з підкріпленням поводить ся максимально близько до того, як відбувається відбір найефективніших рішень у нашій екосистемі. Це дозволяє максимально ефективно перейти від просто випадкового набору атрибутів до такого, що найефективніше взаємодіє із середовищем.

З іншого боку, маємо і недоліки, пов'язані із тим же, що можемо вважати перевагами: а саме те, що велика частина початкових дій є безглуздими.

Як показує стаття, професійним підходом буде ефективно підібрати:

- Ваги
- Функцію винагороди

Це є власне те, що робитиме такі моделі ефективними.

Ключові інсайди

Перспективи навчання без додаткових даних

Навчання з підкріпленням дозволяє навчати у випадках, коли ми не маємо історичних даних, базуючись на результатах фактичних дій агента. Наприклад,

бойові дрони ЗСУ можуть навчати модель, котра покликана навчати та покращувати прошивку їх для більш ефективних дій проти агресора, базуючись на фактичних результатах уже відбутого бойового вильоту. Таких прикладів ефективного використання навчання з підкріпленням можна навести значно більше.

«Природний» відбір найефективніших агентів

Цей принцип приємно вразив мене. Адже дозволяє «закріплювати» найкращу поведінку, просто відсікаючи кожного агента з низькою винагородою, що означає слабку взаємодію агента із середовищем. В результаті навчання за таким методом ми можемо досягнути найефективнішого використання середовища для досягнення наших цілей.

Висновок

У статті було розглянуто принцип роботи навчання з підкріпленням з використанням генетичного алгоритму. Було описано принцип роботи самого алгоритму, а також схему нейромережі, котра реалізовувала навчання з підкріпленням. Якби ще б стаття містила приклади коду, це був би неоціненний посібник із побудови такої нейромережі.