



Laporan

Reinforcement Learning :: Pertemuan 1

Mochammad Dozen K

Banda Subagja

Mohd. Agus Fadillah Sani

Muhammad Alfian Irsyadi Hutagalung

ABD HAMID AFANDI

Better

Atlas

Visioner

Persevere

Alan Turing

03/22/2022

Daftar Isi

Daftar Isi	1
Pengantar	2
Badan	3
Heading 2	3
Heading 3	3
Kesimpulan	4
Rekomendasi	5
Daftar Referensi	6
Lampiran	7

Pengantar

Reinforcement Learning (RL) [in-rID: Pembelajaran Penguatan] adalah area pembelajaran mesin yang berkaitan dengan bagaimana agen cerdas harus mengambil tindakan di lingkungan untuk memaksimalkan gagasan penghargaan kumulatif. Pembelajaran penguatan adalah salah satu dari tiga paradigma pembelajaran mesin dasar, di samping pembelajaran terawasi dan pembelajaran tanpa pengawasan.

Pembelajaran penguatan berbeda dari pembelajaran terawasi dalam hal tidak perlunya pasangan input/output berlabel disajikan, dan dalam tidak membutuhkan tindakan sub-optimal untuk dikoreksi secara eksplisit. Sebaliknya fokusnya adalah menemukan keseimbangan antara eksplorasi (wilayah yang belum dipetakan) dan eksploitasi (pengetahuan saat ini). Algoritma RL yang diawasi sebagian dapat menggabungkan keunggulan algoritma yang diawasi dan RL.

Lingkungan biasanya dinyatakan dalam bentuk Markov Decision Process (MDP), karena banyak algoritma pembelajaran penguatan untuk konteks ini menggunakan teknik pemrograman dinamis. Perbedaan utama antara metode pemrograman dinamis klasik dan algoritma pembelajaran penguatan adalah bahwa yang terakhir tidak mengasumsikan pengetahuan tentang model matematika yang tepat dari MDP dan mereka menargetkan MDP besar di mana metode yang tepat menjadi tidak layak.

Badan

Instruksi

Aktivitas Kelas RL Pertemuan pertama:

1. Cari dan tuliskan satu contoh aplikasi/implementasi RL dibidang anda!
2. Jelaskan secara singkat mengapa aplikasi tersebut merupakan aplikasi berbasis RL, jelaskan berdasarkan kesesuaian dengan karakteristik RL!
3. Dari aplikasi tersebut, tentukan: (a) Objective (b) State (c) Action (d) Reward (e) Termination.
4. Berdasarkan 3 nomor diatas, utarakan/jelaskan pemahaman anda tentang RL, terkait karakteristik, derajat kemudahan/kesulitan untuk diaplikasikan menyelesaikan masalah riil dst.

Jawaban

1. AlphaZero yang dikembangkan untuk menguasai permainan catur.
2. AI AlphaZero menggunakan algoritma *Monte Carlo Tree Search* untuk menentukan gerakan yang akan diambil berdasarkan kecerdasan yang didapat sebelumnya melalui proses machine learning, lebih detailnya melalui metode deep learning dan pelatihan yang berulang ulang.
3. Pengaplikasian AlphaZero terhadap metode Markov Decision Process sebagai berikut:
 - a. Objective: Periksa langkah dan skakmat untuk mengakhiri permainan di mana menang/seri.
 - b. State: .Menunggu lawan bergerak (Giliran), Melindungi bidak kita , atau memakan bidak lawan, mendekati skakmat atau bertahan.
 - c. Action: Pergerakan bidak.
 - d. Reward: (+)Memakan bidak musuh ,(-) Membuat bidak dapat termakan
 - e. Termination: Ketika keadaan pergerakan bidak raja mutlak mati/skakmat.

-
4. Mengembangkan suatu model program komputer yang mana dibuat untuk memecahkan masalah tertentu*.

RL membutuhkan daya komputasi yang besar jika digunakan untuk memecahkan masalah yang rumit, dikaitkan dengan konteks kita menghadapi masalah di real life, mungkin masalah yang kita hadapi bisa sederhana, tetapi mungkin kita akan menggunakan RL untuk beberapa masalah yang tidak bisa dilakukan oleh daya manusia dengan metode manual, karena kita tidak bisa langsung menggunakan AI dan mengganti semua pekerja manusia.

Selain itu AI meskipun sangat powerful, dan prospeknya sangat menjanjikan, tetapi membutuhkan pengembangan yang sangat lama agar bekerja dengan baik sampai bisa diandalkan untuk menangani suatu pekerjaan dengan baik. Metode RL membutuhkan banyak waktu untuk mengembangkannya sampai bisa diimplementasikan di real life, selain itu untuk implementasi real life mungkin membutuhkan fasilitas/ environment simulasi/training/ learning yang bisa me replika environment real life, dan mungkin itu mahal.

Kesimpulan

.

Rekomendasi

.

Daftar Referensi

1. Kaelbling, Leslie P.; Littman, Michael L.; Moore, Andrew W. (1996). "Pembelajaran Penguatan: Sebuah Survei". *Jurnal Penelitian Kecerdasan Buatan*. Jil. 4 hlm. 237–285.
arXiv: [\[cs/9605103\] Reinforcement Learning: A Survey](https://arxiv.org/abs/cs/9605103) [arxiv.org]
doi: [Reinforcement Learning: A Survey](https://doi.org/10.1006/jair.1996.0026) [Journal of Artificial Intelligence Research]
S2CID: [\[PDF\] Reinforcement Learning: A Survey](https://www.semanticscholar.org/paper/Reinforcement-Learning-A-Survey-Littman-Kaelbling-Moore/10.1006/jair.1996.0026) [Semantic Scholar]
2. Pandian, B. Jaganatha; Noel, Mathew Mithra (2018-09-01). "Kontrol bioreaktor menggunakan algoritma pembelajaran penguatan baru yang diawasi sebagian". *Jurnal Pengendalian Proses*. Jil. 69 hlm. 16–29.
doi: [Control of a bioreactor using a new partially supervised reinforcement learning algorithm](https://doi.org/10.1016/j.jprocont.2018.09.001) [Science Direct]
ISSN: [Journal of process control \(eJournal / eMagazine, 1995\)](https://www.worldcat.org/issn/0959-1526) [WorldCat.org]
S2CID: [Control of a bioreactor using a new partially supervised reinforcement learning algorithm](https://www.semanticscholar.org/paper/Control-of-a-bioreactor-using-a-new-partially-supervised-reinforcement-learning-algorithm-Pandian-Noel/10.1016/j.jprocont.2018.09.001) [Semantics Scholar]
3. van Otterlo, M.; Wiering, M. (2012). *Penguatan pembelajaran dan proses keputusan Markov. Pembelajaran Penguatan*. Adaptasi, Pembelajaran, dan Optimalisasi. Jil. 12. hlm. 3-42.
doi: [Reinforcement Learning and Markov Decision Processes](https://doi.org/10.1007/978-3-642-27644-6) [SpringerLink]
ISBN: 978-3-642-27644-6 [[Google Books](https://books.google.com/books?id=9783642276446) | [Open Library](https://openlibrary.org/books/OL27644618M) | [Amazon](https://www.amazon.com/Reinforcement-Learning-Markov-Decision-Processes/dp/9783642276446)]

Lampiran

Tautan tangkapan layar: .