## Ứng dụng : Điều khiển đèn giao thông

Với sự gia tăng đô thị hóa và sự gia tăng số lượng ô tô mỗi hộ gia đình, tắc nghẽn giao thông đã trở thành một vấn đề lớn, đặc biệt là ở các khu vực đô thị.

Học tăng cường là một phương pháp tiếp cận dựa trên dữ liệu theo xu hướng để điều khiển tín hiệu giao thông thích ứng. Các mô hình này được đào tạo với mục tiêu học một chính sách sử dụng hàm giá trị điều khiển đèn giao thông một cách tối ưu dựa trên tình trạng giao thông hiện tại.

Việc ra quyết định cần linh hoạt tùy thuộc vào tốc độ đến của lưu lượng truy cập từ các hướng khác nhau, tỷ lệ này phải thay đổi vào các thời điểm khác nhau trong ngày. Cách xử lý lưu lượng thông thường dường như bị hạn chế do hành vi không cố định này. Ngoài ra, không thể sử dụng lại   chính sách *π được đào tạo cho giao lộ có x làn đường ở giao lộ có y làn đường.*

Diagram

Description automatically generated

Học tăng cường (RL) là một cách tiếp cận thịnh hành do tính chất dựa trên dữ liệu của nó để điều khiển tín hiệu giao thông thích ứng trong các mạng giao thông đô thị phức tạp.

Có một số hạn chế trong việc áp dụng các thuật toán Học tăng cường sâu cho mạng lưới giao thông, chẳng hạn như tình thế tiến thoái lưỡng nan trong thăm dò-khai thác, kế hoạch đào tạo đa tác nhân, không gian hành động liên tục, phối hợp tín hiệu, v.v.



### **phẫu thuật robot**

Một ứng dụng Học tăng cường mạnh mẽ trong quá trình ra quyết định là việc sử dụng robot phẫu thuật có thể giảm thiểu lỗi và bất kỳ biến thể nào và cuối cùng sẽ giúp nâng cao hiệu quả của bác sĩ phẫu thuật. Một trong những robot như vậy là Da Vinci , cho phép các bác sĩ phẫu thuật thực hiện các quy trình phức tạp với sự linh hoạt và khả năng kiểm soát cao hơn so với các phương pháp thông thường.

Các tính năng quan trọng được cung cấp là hỗ trợ bác sĩ phẫu thuật bằng các dụng cụ tiên tiến, dịch chuyển động tay của bác sĩ phẫu thuật trong thời gian thực và cung cấp chế độ xem 3D độ nét cao về khu vực phẫu thuật.

## Gaming

Học tăng cường đã tiếp quản các phương pháp tạo trò chơi điện tử truyền thống.

So với các trò chơi điện tử truyền thống, nơi chúng ta cần có một cây hành vi phức tạp để tạo ra logic của trò chơi, thì việc đào tạo một mô hình Học tăng cường đơn giản hơn nhiều. Tại đây, tác nhân được thiết lập để tự học trong môi trường trò chơi mô phỏng bằng cách thực hiện chuỗi hành động cần thiết để đạt được hành vi mong muốn.

Trong Học tăng cường, đặc vụ phải được đào tạo về tất cả các khía cạnh của trò chơi như tìm đường, phòng thủ, tấn công và tạo các chiến lược dựa trên tình huống để khiến trò chơi trở nên thú vị đối với đối thủ.

Tùy thuộc vào trí thông minh mà bot thu được, các cấp độ của trò chơi được thiết lập.



Google DeepMind là một ví dụ sống động về Tối ưu hóa trò chơi.

Chúng ta đã thấy trong AlphaGo, một đặc vụ được đào tạo bởi RL đã đánh bại kỳ thủ cờ vây mạnh nhất trong lịch sử ghi một bàn thắng được coi là không thể vào thời điểm đó. Nó được biết đến là một trò chơi rất thử thách dành cho Trí tuệ nhân tạo.

AlphaGo, một chương trình máy tính do DeepMind, một công ty của Google, tạo ra, sử dụng sự kết hợp của cây tìm kiếm nâng cao và mạng lưới thần kinh sâu. Các mạng nơ-ron này lấy bàn cờ vây làm đầu vào dẫn xuất các tính năng thông qua các lớp mạng khác nhau chứa hàng triệu kết nối giống như nơ-ron.

Các tác nhân Học tăng cường cũng được sử dụng trong phát hiện lỗi và thử nghiệm trò chơi. Điều này là do khả năng chạy rất nhiều lần lặp mà không cần đầu vào của con người, kiểm tra căng thẳng và tạo tình huống cho các lỗi tiềm ẩn.

Các công ty trò chơi mới hơn như Ubisoft gần đây đã sử dụng Học tăng cường để giảm số lượng lỗi đang hoạt động được tìm thấy trong trò chơi. Các đặc vụ RL được đào tạo trong môi trường trò chơi bằng cách sử dụng các kỹ thuật thăm dò và khai thác để kiểm tra một số cơ chế trò chơi nhằm cố gắng khắc phục chúng.

## ô tô tự lái

Việc lái xe trong môi trường bối cảnh mở phải được hỗ trợ bởi [mô hình máy học](https://www.v7labs.com/blog/machine-learning-guide) được đào tạo với tất cả các cảnh và tình huống có thể xảy ra trong thế giới thực.

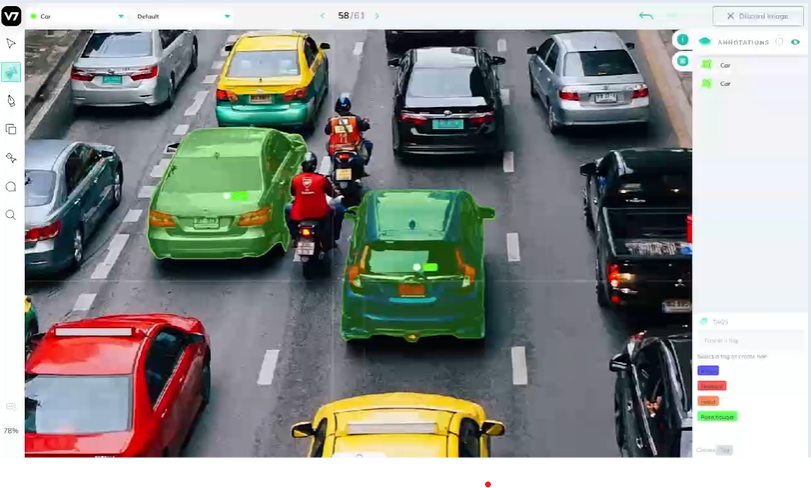
Tuy nhiên-

Việc tập hợp các loại cảnh này là một vấn đề phức tạp cần giải quyết. Làm thế nào chúng ta có thể đảm bảo rằng một chiếc xe tự lái đã học được tất cả các tình huống có thể xảy ra và làm chủ mọi tình huống một cách an toàn?

Câu trả lời cho điều này là Học tăng cường .

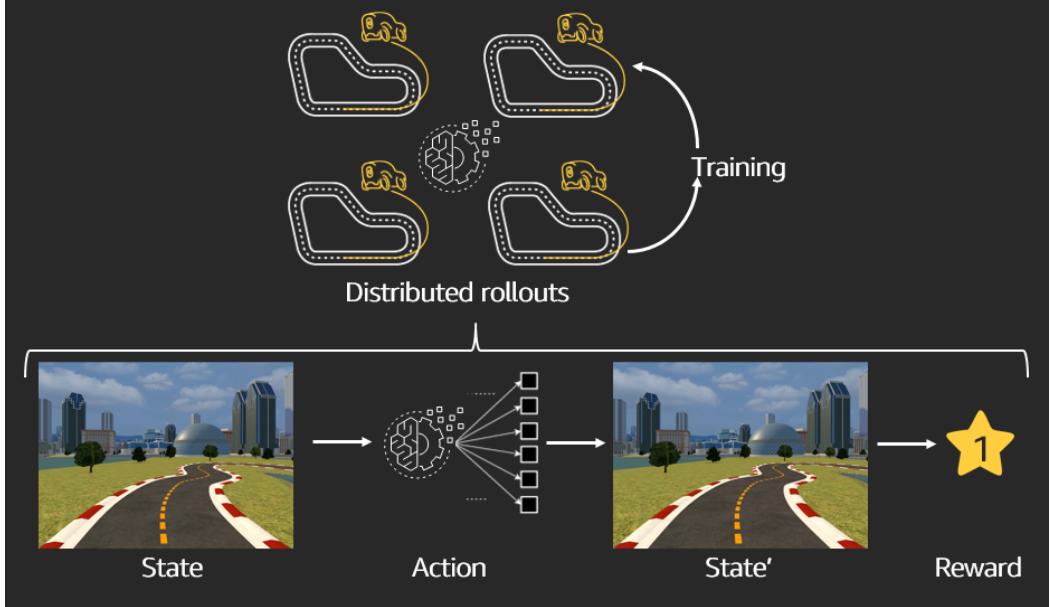
Các mô hình Học tăng cường được đào tạo trong một môi trường năng động bằng cách học một chính sách từ kinh nghiệm của chính nó tuân theo các nguyên tắc thăm dò và khai thác nhằm giảm thiểu sự gián đoạn đối với lưu lượng truy cập. Xe tự lái có nhiều khía cạnh để xem xét tùy theo đó nó đưa ra quyết định tối ưu.

Khu lái xe, xử lý giao thông, duy trì giới hạn tốc độ, tránh va chạm là những yếu tố quan trọng.



Nhiều môi trường mô phỏng có sẵn để thử nghiệm các mô hình Học tăng cường cho các công nghệ xe tự lái.

DeepTraffic là một môi trường mã nguồn mở kết hợp sức mạnh của Học tăng cường, [Học sâu](https://www.v7labs.com/blog/deep-learning-guide) và [Thị giác máy tính](https://www.v7labs.com/blog/what-is-computer-vision) để xây dựng các thuật toán được sử dụng cho lái xe tự động do MIT đưa ra. Nó mô phỏng các phương tiện tự trị như máy bay không người lái, ô tô, v.v.



Lái xe tự hành sử dụng Học tăng cường với sự trợ giúp của các môi trường tổng hợp này để giải quyết các vấn đề quan trọng về **tối ưu hóa quỹ đạo** và **đường dẫn động.**

Các tác nhân Học tăng cường được đào tạo trong các môi trường năng động này để tối ưu hóa quỹ đạo. Các tác nhân học lập kế hoạch chuyển động, thay đổi tuyến đường, quyết định và vị trí đỗ xe và kiểm soát tốc độ, v.v.

**Ưu điểm reinforcement learning:**

* Reinforcement learning có thể giải quyết các vấn đề rất phức tạp mà không thể được giải quyết bằng các kỹ thuật thông thường.
* Kỹ thuật này được ưa chuộng để đạt được các kết quả lâu dài, rất khó đạt được.
* Mô hình học này rất giống với quá trình học của con người. Do đó, nó gần đến việc đạt tới sự hoàn hảo.
* Mô hình có thể sửa chữa các lỗi xảy ra trong quá trình huấn luyện.
* Một khi một lỗi được sửa chữa bởi mô hình, khả năng xảy ra lỗi tương tự là rất ít.
* Nó có thể tạo ra mô hình hoàn hảo để giải quyết một vấn đề cụ thể.
* Các robot có thể triển khai các thuật toán học tăng cường để học cách đi bộ.
* Trong trường hợp không có tập dữ liệu huấn luyện, nó bị ràng buộc phải học từ kinh nghiệm của nó.
* Các mô hình học tăng cường có thể vượt trội hơn con người trong nhiều nhiệm vụ. Chương trình AlphaGo của DeepMind, một mô hình học tăng cường, đã đánh bại nhà vô địch thế giới Lee Sedol trong trò chơi Go vào tháng 3 năm 2016.
* Reinforcement learning được thiết kế để đạt được hành vi lý tưởng của một mô hình trong ngữ cảnh cụ thể để tối đa hóa hiệu suất.
* Nó có thể hữu ích khi cách duy nhất để thu thập thông tin về môi trường là tương tác với nó.
* Thuật toán học tăng cường duy trì sự cân bằng giữa khám phá và khai thác. Khám phá là quá trình thử nghiệm những thứ khác nhau để xem chúng có tốt hơn những gì đã được thử trước đó hay không. Khai thác là quá trình thử những thứ đã hoạt động tốt nhất trong quá khứ. Các thuật toán học khác không thực hiện được sự cân bằng này.

**Nhược điểm reinforcement learning:**

1. Yêu cầu lượng dữ liệu lớn và tính toán phức tạp: Reinforcement learning yêu cầu lượng dữ liệu lớn và tính toán phức tạp để học được chính xác hành vi tối ưu. Điều này có nghĩa là mô hình học sâu và hệ thống phần cứng tính toán phải được cấu hình để đáp ứng yêu cầu này.
2. Cần phải giải quyết vấn đề độc lập Markov: Phương pháp này giả định rằng trạng thái hiện tại chỉ phụ thuộc vào trạng thái trước đó, mà không cần quan tâm đến lịch sử các trạng thái trước đó. Tuy nhiên, trong thực tế, rất nhiều bài toán phức tạp yêu cầu thông tin về lịch sử để đưa ra quyết định tối ưu.
3. Khó khăn khi ứng dụng trong các bài toán đơn giản: Khi sử dụng phương pháp này để giải quyết các bài toán đơn giản, có thể sẽ gặp phải vấn đề sử dụng công cụ quá phức tạp và tốn kém để giải quyết các bài toán đơn giản này.
4. Khó khăn khi sử dụng trong các hệ thống thực tế: Reinforcement learning có thể gặp khó khăn khi áp dụng vào các hệ thống thực tế vì các hệ thống này thường rất phức tạp và các giải pháp tối ưu của reinforcement learning có thể không phù hợp cho các vấn đề thực tế.
5. Chi phí để xây dựng và bảo trì các hệ thống robot và các thiết bị liên quan: Các thiết bị liên quan đến robot và các hệ thống tương tự thường rất đắt tiền, dễ bị hỏng và yêu cầu bảo trì cẩn thận. Việc sửa chữa và bảo trì các hệ thống này có thể rất tốn kém và tốn nhiều thời gian.