

- ۱) یک موضوع برای کار تحقیقاتی خود پستیها در دهید.
- ۲) این پروژه چه اهدافی را دنبال میکند؟ (کدامانی، کلی، ویژه، کاربردی)
- ۳) این تحقیق چه مشکلی را می‌خواهد حل کند؟
- ۴) یک دیاگرام استخوان ماهی برای این طرح رسم کنید.

موضوع: یادگیری تقویتی در روباتیک

Reinforcement learning in Robotics

یادگیری تقویتی (Reinforcement learning) یکی از زیرشاخه‌های machine learning (یادگیری ماشین) است که از روانشناسی زنبان برای الهام می‌گیرد. این روش بر رفتارهایی تمرکز می‌کند که ماشین باید برای حداکثر کردن پاداشش انجام دهد. این علم با تجربه به گسترده بودن، در زمینه‌های مختلفی چون: نظریه بازی‌ها، سیستم‌های از راه دور، آمار، الگوریتم ژنتیک و... کاربرد دارد.

تفاوت اصلی این روش، با روش‌های سنتی یادگیری ماشین آن است که <sup>بازاف</sup> آن روش‌های (supervised و unsupervised)، ماشین ما نیازی به دانستن اطلاعات دنیای

فرکانس تعیین نمی‌کند.

همین، این نوع یادگیری، دو نوع تفاوت عمده با روش‌های قبلی دارد. نخست آن که در آن، نوع داده ورودی و خروجی (نوع صبح) وجود ندارد و رفتارها از بیرون اصلاح نمی‌شوند. دوم آن که تمرکز بر روی کارایی زنبان وجود دارد که نیازمند یاد کردن یک تکامل مناسب بین کشف رفتار جدید و بهینه‌سازی از آن‌ها دارد.

نکته: اینها:

- S: یک مجموعه از حالات مختلف مسئله
- A: یک مجموعه از اقدامات
- R: مجموعه‌ای قوای گذار
- RG: فعلین اهداف یا پاداش به ازای هر وضعیت

نکته مهم در این بابست که یادگیری تقویتی برای مسائلی که در آن‌ها نتیجه سود در کوتاه مدت، تعیین کننده ی نتیجه سود در درازمدت نیست، بسیار مناسب است.

چرا که در این روش، نکته کلیدی آن است که ماشین، در هر مرحله، لزوماً بهترین راه را انتخاب نمی کند و در نهایت هم سعی این برای این است که مجموع پاداش را به حداکثر برساند. که این ویژگی باعث می شود، که این روش در کنترل ربات ها، برنامه ریزی آگناسیوها، مدیریت ترافیک، تخصیص منابع و... به کار رود.

عوامل برتری این روش: ۱) استفاده از نمونه ها به جای مدل سازی گرایه (۲) استفاده از چند منابع

program (A) برای مساله یک مدل وجود دارد

(B) فقط یک مدل نتیجه سودی است که برای مدل در دسترس است

(C) تنها راه برای به دست آوردن اطلاعات از محیط، تعامل با آن است (ناظر نداشت)

↓  
یادگیری از تجربه

هر سه روش، در قالب یادگیری تقویتی، تبدیل به یک مساله یادگیری می شوند.

قلب کلی این نوع یادگیری به این صورت است که یک سری مأمور (Agent) داریم و این مأموران وظایف دارند با محیط اطرافشان تعامل داشته باشند و یاد بگیرند که با پاسخهای عددی را به حداکثر برسانند. در واقع اثر انگشت کنند، آن‌ها را تئیه می کنند و اثر درست عمل کنند، به آن ها پاداش می دهند.

یادگیری مکتوب:

هیچ تابع پاداشی وجود ندارد، روبات سعی می کند که با مشاهده یک رفتار که معرکه به رفتار کینه ترنیک است، آن را تقلید کند.



اهداف پروژه:

آزمایی

هدف یادگیری ماشین این است که کامپیوتر (در کمترین مفهوم خود) بتواند به تدریج و با افزایش داده های بازدهی بالاتری در وظیفه مورد نظر خود پیدا کند. گسترده ای این وظیفه می تواند از تشخیص عواید هرچه با دیدن چند نمونه از چهره مورد نظر تا فراگیری شیوه های کاربردی برای رویت های خاص (فوتبال و غیره) با دریافت سیگنال پاداش و تنبیه باشند.

کلیدی:

طبقه بندی های این زمینه تقریباً گسترده است.

از یک سو برآنند که روش های یادگیری ماشینی به وجود بینورند (هدف ویژه ۱) و امکان پذیری و کیفیت یادگیری را برای روش هایشان مطالعه کنند (هدف ویژه ۲) عده ای از پژوهشگران سعی دارند روش های یادگیری ماشینی را به مسائل تازه ای اقبال کنند (هدف ویژه ۳)

البته اکثر مسائل، معمولاً "دو یا سه رویکرد را دنبال می کنند و در یک طبقه قابل انتخابی نیستند.

کاربرد:

دکتر ج. کوب و ج. اندرو بگنل و ج. پتند در مقاله ای در این باب می نویسند:

"یادگیری تقویتی یک سری وسایط و سبته مناسب برای رویت ها فراهم می کند، که برای آموزش روش های مهندسی شده / رفتارهای خاص به آن ها مناسب است."

از سوی دیگر، می توان دیدن کاربردی در این حوزه در نظر گرفت.

یکی اینکه آن چه برای انسان پیچیده است، می تواند برای یک ربات بسیار ساده باشد. بخش این موضوع هم حلقه است.

در نهایت هدف کاربردی یادگیری تقویتی در روباتیک بدین صورت است که ماشین ما به صورتی ماهر شود که با حساسیت کافی کار کند و آمادگی پیگیری وظایف مختلف از سمت محیط را داشته باشد.

در سطوح بالاتر کاربرد، گاهی نیاز است روش‌های خاصی به کار برده شود تا محیط اصلاح شود که بدین ترتیب روبات بتواند بهتر عمل کند.

تحلیل در سطوح مارکیٹینگ و بازاریابی می‌تواند از Agentها (ماموران) برای نتایج سطح رضایت مشتریان و میزان فروش و ... استفاده کرد.