

دانشگاه صنعتی امیرکبیر  
 (پلی‌تکنیک تهران)

دانشکده مهندسی کامپیوتر

گزارش درس روش پژوهش و ارائه

یادگیری تقویتی عمیق در سیستم‌های توصیه‌گر

نگارش

آرین قزوینی

استاد راهنما

دکتر مهدی صدیقی

اردیبهشت ۱۴۰۲

# چکیده

ذکر کاربردهای محبوب مورداستفاده امروزه از سیستم‌های توصیه‌گر مانند کاربرد آن‌ها در توصیه درمان‌های پزشکی، استفاده در شبکه‌های اجتماعی، سیستم پخش فیلم و...

دلایل رشد محبوبیت یادگیری تقویتی عمیق

اشاره مختصر به تفاوت بین سیستم‌های توصیه‌گر با روش‌های بدون یادگیری و استفاده از یادگیری تقویتی در آنها

|  |  |
| --- | --- |
| فهرست مطالب (ساختار اولیه) | صفحه |

[فصل اول : مقدمه](#_Toc28770795)

[فصل دوم: مروری بر یادگیری ماشین و شبکه های عصبی](#_Toc28770796)

۲-یادگیری ماشین..........................................................................................................................................................................

۱-۲-یادگیری ماشین چیست.......................................................................................................................................................  
۲-۲-اهمیت یادگیری ماشین.......................................................................................................................................................  
۳-۲-انواع سیستم های یادگیری ماشین....................................................................................................................................  
 ۱-۳-۲-یادگیری با نظارت........................................................................................................................................................  
 ۲-۳-۲-یادگیری بدون نظارت.................................................................................................................................................  
 ۳-۳-۲-یادگیری تقویتی...........................................................................................................................................................  
 ۴-۳-۲-یادگیری دسته‌ای..........................................................................................................................................................  
۵-۳-۲ -یادگیری برخط................................................................................................................................................................  
۴-۲-شبکه های عصبی و یادگیری عمیق..................................................................................................................................

۴-۲-خلاصه.....................................................................................................................................................................................

فصل سوم: مفاهیم اولیه یادگیری تقویتی.......................................................................................

۳-یادگیری تقویتی.................................................................................................................................................

۳-۱-فرآیند تصمیم گیری مارکوف.......................................................................................................................  
۳-۲-انواع الگوریتم های یادگیری تقویتی............................................................................................  
۳-۳-یادگیری تقویتی بدون شبکه های عصبی.............................................................................................

۳-۳-۱-روش مونت کارلو و تفاضل زمانی......................................................................................  
۲-۳-۳-روش یادگیری Q...............................................................................................................

۴-۳-یادگیری تقویتی با شبکه های عصبی............................................................................................................

۱-۴-۳-روش شبکه های Qعمیق..........................................................................................................  
۲-۴-۳-روش بازیگر منتقد......................................................................................................................

۵-۳-خلاصه...................................................................................................................................................................

[فصل چهارم: یادگیری تقویتی در سیستم های توصیه‌گر](#_Toc28770816)

[فصل پنجم: نتیجه گیری و پیشنهادات](#_Toc28770830)

[۱-۵-نتیجه گیری](#_Toc28770831)

۲-۵-پیشنهادات...........................................................................................................................................................................

**منابع**....................................................................................................................................................................................

**فصل دوم  
مروری بر یادگیری ماشین و شبکه های عصبی**

**۲- یادگیری ماشین**

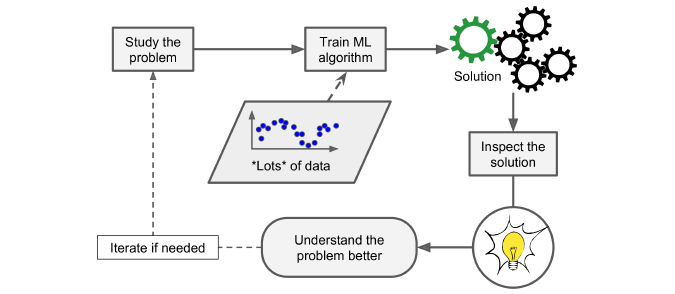
در این بخش ابتدا به مفاهیم اولیه یادگیری ماشین پرداخته و سپس به اهمیت استفاده از یادگیری ماشین در جهان امروز، انواع الگوریتم‌های یادگیری ماشین و در نهایت به مفاهیم کلی شبکه‌های عصبی و یادگیری عمیق اشاره می‌شود.

**۱-۲- یادگیری ماشین چیست؟**

یادگیری ماشینی شامل برنامه نویسی رایانه ها به گونه ای است که آنها می توانند دانش را از داده ها به دست آورند.فرآیند یادگیری در یادگیری ماشین زمانی اتفاق می‌افتد که یک برنامه عملکرد خود را در یک کار خاص با استفاده از تجربیات یا داده‌هایی که با معیار عملکرد اندازه‌گیری می‌شوند، افزایش می‌دهد. برای نشان دادن این مفهوم، می‌توان به فیلتر هرزنامه[[1]](#footnote-1) اشاره کرد که شناسایی ایمیل ها را بر اساس نمونه هایی از ایمیل های هرزنامه و غیر هرزنامه یاد می گیرد. مجموعه مثال‌هایی که برای یادگیری استفاده می‌شوند، مجموعه آموزشی نامیده می‌شوند، که هر مثال را یک نمونه آموزشی می‌نامند.در این سناریو، معیار عملکرد می‌تواند تعداد ایمیل هایی باشد که به صورت صحیح تفکیک شده‌اند.شایان ذکر است که به‌دست آوردن حجم وسیعی از داده‌ها، مانند دانلود ویکی‌پدیا، ذاتاً توانایی‌های رایانه را در هیچ کار خاصی بهبود نمی‌بخشد و در نتیجه واجد شرایط یادگیری ماشینی نمی‌شود]1[.

**۲-۲- اهمیت یادگیری ماشین**

یادگیری ماشینی می تواند یادگیری انسان را تسهیل کند. به عنوان مثال ، هنگامی که یک فیلتر هرزنامه با داده‌های کافی آموزش داده شده باشد، می توان آن را بررسی کرد تا کلمات و ترکیبات کلماتی را که به عنوان پر تکرارترین شاخص های اسپم در نظر می گیرد، آشکار کند. این فرآیند می‌تواند الگوهای دیده نشده در داده‌ها را مشخص کند و در نتیجه به درک بهتر مساله کمک می‌کند]1[.شکل ۲ فرآیند بهبود درک مساله با استفاده از یادگیری ماشین نشان می‌دهد.



**شکل۲-یادگیری ماشین برای درک بهتر مساله**]1[

**۳-۲- انواع سیستم‌های یادگیری ماشین**

با توجه به طیف گسترده‌ای از سیستم‌های یادگیری ماشینی موجود، دسته‌بندی آنها بر اساس معیار‌های مختلفی می‌تواند صورت بگیرد:

* نظارت: این دسته بندی در نظر می گیرد که آیا سیستم‌ها تحت نظارت انسان آموزش داده می‌شوند یا خیر. این شامل یادگیری تحت نظارت[[2]](#footnote-2)، یادگیری بدون نظارت[[3]](#footnote-3)، یادگیری نیمه نظارتی[[4]](#footnote-4)، و یادگیری تقویتی [[5]](#footnote-5)است.
* یادگیری افزایشی: این طبقه بندی تشخیص می دهد که آیا سیستم‌ها توانایی یادگیری تدریجی با رسیدن داده‌های جدید را دارند یا خیر. از الگوریتم های معروف آن می‌توان به یادگیری آنلاین [[6]](#footnote-6)و یادگیری دسته‌ای [[7]](#footnote-7)اشاره نمود.
* رویکرد یادگیری: این دسته بندی نحوه عملکرد سیستم ها را هنگام مواجهه با داده های جدید بررسی می کند. این شامل دو رویکرد اصلی است: یادگیری مبتنی بر نمونه، که در آن نقاط داده جدید با داده های موجود مقایسه می شود، و یادگیری مبتنی بر مدل، که در آن الگوهای موجود در داده های آموزشی برای ساخت یک مدل بکار می‌رود.

**۱-۳-۲- یادگیری با نظارت**

یادگیری نظارت شده، روشی است که بر مجموعه داده های برچسب گذاری شده برای آموزش و پیش بینی نتیجه تکیه می کند. در طول فرآیند تمرین، مدل وزن های خود را برای دستیابی به تناسب مناسب تنظیم می کند.از کاربردهای معروف یادگیری تحت نظارت می‌توان به پیش‌بینی مقادیر عددی مانند قیمت خانه براساس ویژگی های داده شده برای هر نمونه از داده‌های آموزشی اشاره کرد که به این تکنیک رگراسیون گفته می شود.یادگیری نظارت شده شامل تکنیک های مختلفی از جمله شبکه های عصبی، بیز ساده، رگرسیون خطی، رگرسیون لجستیک، جنگل تصادفی و ماشین بردار پشتیبان (SVM) است, 2]1[.

**۲-۳-۲- یادگیری بدون نظارت**

# مقدمه

۱-بیان انگیزه شکل‌گیری سیستم‌های توصیه‌گر

۲-تاریخچه مختصر از سیستم‌های توصیه‌گر

۳-اهمیت سیستم‌های توصیه‌گر در شرکت‌های بزرگ تکنولوژی

۴-معرفی کلی الگوریتم‌های بدون یادگیری همچون: فیلتر مشارکتی، فیلتر مبتنی بر محتوا، فیلتر هیبریدی

۵-بیان مختصر مشکلات ۳ الگوریتم مطرح شده

۶-ویژگی های مسائلی که می‌توانند با یادگیری تقویتی به‌خوبی حل شوند.

**بند اصلی**

تعریف مفهوم یادگیری ماشین  
ذکر چند مثال از کاربردهای امروزه یادگیری ماشین  
مقایسه روش‌های سنتی تشخیص ایمیل‌های اسپم از ایمیل‌های غیر اسپم با کاربرد یادگیری ماشین در این حوزه  
بیان ویژگی مسائلی که به‌خوبی با روش‌های یادگیری ماشین حل می‌شوند  
اشاره کلی به روش‌های دسته‌بندی سیستم‌های یادگیری ماشین  
توضیح ایده کلی یادگیری با نظارت   
ذکر چند الگوریتم معروف یادگیری با نظارت  
توضیح ایده کلی یادگیری بدون نظارت  
ذکر مثال‌های معروف الگوریتم‌های یادگیری بدون نظارت  
شرح ایده یادگیری نیمه نظارتی   
ذکر مثال نسبت‌دادن اسم افراد به تصاویر آنها برای یادگیری نیمه نظارتی  
اشاره به ایده کلی یادگیری دسته‌ای   
اشاره به ایده کلی یادگیری برخط  
شرح ایده کلی یادگیری تقویتی  
معرفی مفاهیم اولیه شبکه‌های عصبی و مدل‌سازی پرسپترون از نورون‌های مغزی  
توضیح فرایند تصمیم‌گیری مارکوف و معادله بلمن  
اشاره مختصر به دسته‌بندی الگوریتم‌های یادگیری تقویتی (تفاوت الگوریتم‌های با مدل و بدون مدل)   
شرح ایده الگوریتم مونت‌کارلو  
معرفی الگوریتم Temporal Difference Learning  
شرح تفاوت‌های MC,TD  
 معرفی روش SARSA  
 معرفی روش Q-learning  
بیان علت استفاده از شبکه‌های عصبی در یادگیری تقویتی   
معرفی روش شبکه عمیق Q  
معرفی کلی روش‌های مبتنی بر سیاست  
 معرفی روش گرادیان سیاست و گرادیان سیاست عمیق  
اشاره کلی به روش بازیگر منتقد  
توضیح مصالحه بین دو مفهوم جستجو و به‌کارگیری سیاست در یادگیری تقویتی  
شرح ایده اپسیلون - حریصانه برای مصالحه شرح داده شده  
معرفی چند مورد از چالش‌های یادگیری تقویتی  
معرفی الگوریتم تکرار تجربه به‌عنوان روشی برای بهبود بهره‌وری از داده‌ها  
 ویژگی‌های یادگیری تقویتی که آن را به روشی سودمند برای سیستم‌های توصیه‌گر تبدیل می‌کند.   
معرفی روش‌هایی برای تبدیل ساختار یک سیستم توصیه‌گر به ساختار مسئله تصمیم‌گیری مارکوف (چگونگی تعریف حالات و پاداش‌های عامل)   
اشاره به روش‌های ارزیابی مدل  
مقایسه چند روش یادگیری تقویتی عمیق در سیستم‌های توصیه‌گر از منظر ۴ عامل بیان شده (تعریف حالت، بهینه‌سازی سیاست، تعریف پاداش یا جریمه برای عامل، نحوه تست و ارزیابی)

# خلاصه

بررسی ایده و کاربرد های سیستم های توصیه گر

مرور مفاهیم اولیه یادگیری تقویتی و بررسی روش های ترکیب آن با شبکه‌های عصبی به منظور بهبود کارکرد آن

ذکر چهارچوبی با ۴ عنصر: نمایش حالات، بهینه‌سازی سیاست، فرمول‌بندی پاداش و ارزیابی محیط، برای ساخت سیستم‌های توصیه‌گر با یادگیری تقویتی

|  |
| --- |
|  |

1. Spam [↑](#footnote-ref-1)
2. Supervised Learning [↑](#footnote-ref-2)
3. Unsupervised Learning [↑](#footnote-ref-3)
4. Semi-Supervised Learning [↑](#footnote-ref-4)
5. Reinforcement Learning [↑](#footnote-ref-5)
6. Online Learning [↑](#footnote-ref-6)
7. Batch Learning [↑](#footnote-ref-7)