انواع الگوریتم‌های مورد استفاده در یادگیری عمیق

یادگیری عمیق چیست؟

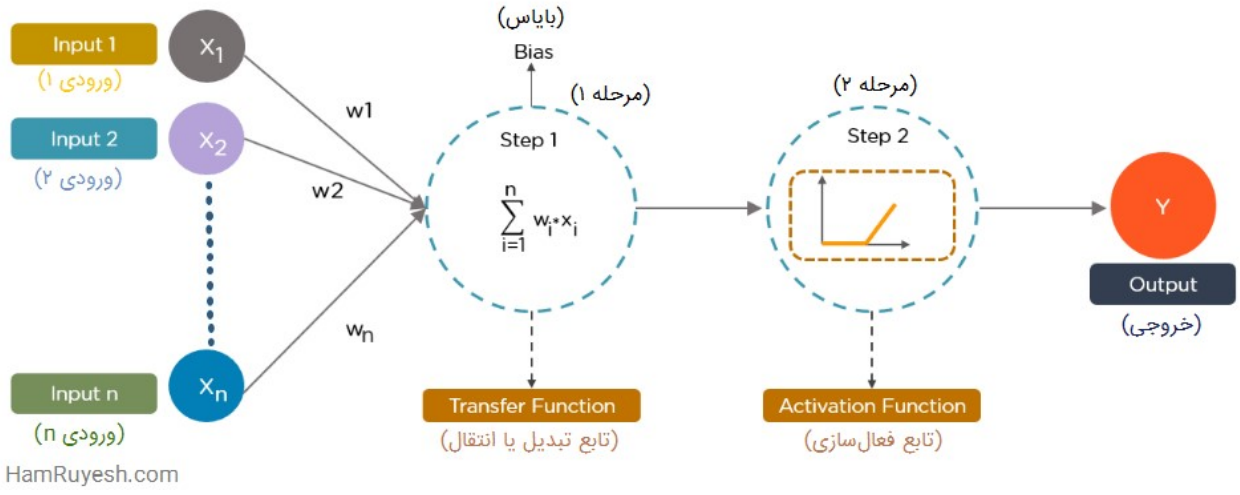
یادگیری عمیق(Deep learning) ، از شبکه‌های عصبی مصنوعی برای انجام محاسبات پیچیده بر روی حجم زیادی از داده‌ها استفاده می‌کند. یادگیری عمیق، نوعی از یادگیری ماشین (machine learning) است که بر اساس ساختار و عملکرد مغز انسان کار می‌کند.

الگوریتم‌های یادگیری عمیق، با یادگیری از مثال‌ها و نمونه‌ها، به ماشین‌ها آموزش می‌دهند. صنایعی مانند مراقبت‌های بهداشتی، تجارت الکترونیک، سرگرمی و تبلیغات، معمولا از یادگیری عمیق استفاده می‌کنند.

تعریف شبکه‌های عصبی

یک شبکه عصبی(neural network) ، مانند مغز انسان ساختار یافته است و از نورون‌های مصنوعی تشکیل شده است که با عنوان "گره" نیز شناخته می‌شود. این گره‌ها در سه لایه در کنار هم چیده شده‌اند:

- لایه ورودی
- لایه (های) مخفی
- لایه خروجی



داده‌ها، اطلاعاتی را در قالب ورودی به هر گره ارائه می‌دهند. گره، ورودی‌ها را در وزن‌های تصادفی ضرب می‌کند، آن‌ها را محاسبه کرده و یک بایاس به آن اضافه می‌کند. در نهایت، توابع غیرخطی، که به آن‌ها "توابع فعال سازی" گفته می شود، برای تعیین اینکه کدام نورون شلیک کند، اعمال می‌شوند.

الگوریتم‌های یادگیری عمیق چگونه کار می‌کنند؟

در حالی که الگوریتم‌های یادگیری عمیق دارای بازنمایی‌های خودآموز (self-learning) هستند، این الگوریتم‌ها به شبکه‌های عصبی مصنوعی وابسته اند که نحوه محاسبه اطلاعات توسط مغز را منعکس می‌کنند. در طول فرایند آموزش، الگوریتم‌ها از عناصر ناشناخته در توزیع ورودی برای استخراج ویژگی‌ها، گروه بندی اشیا و کشف الگوهای مفید داده‌ها استفاده می‌کنند. مشابه آموزش ماشین‌ها برای خودآموزی، این امر در سطوح مختلف، با استفاده از الگوریتم‌هایی برای ساخت مدل‌ها رخ می‌دهد.

مدل‌های یادگیری عمیق از چندین الگوریتم استفاده می‌کنند. با اینکه هیچ شبکه‌ای کامل تلقی نمی‌شود، اما برخی از الگوریتم‌ها برای انجام کارهای خاص مناسب تر هستند. برای انتخاب الگوریتم‌های مناسب، خوب است که یک درک کامل از تمام الگوریتم‌های اولیه داشته باشیم.

انواع الگوریتم‌های مورد استفاده در یادگیری عمیق

در ادامه، ۱۰ مورد از محبوب ترین الگوریتم‌های یادگیری عمیق را مشاهده می‌کنید:

۱. شبکه‌های عصبی کانولوشنی(Convolutional Neural Networks)
۲. شبکه‌های حافظه طولانی کوتاه مدت(Long Short Term Memory Networks)
۳. شبکه‌های عصبی بازگشتی(Recurrent Neural Networks)
٤. شبکه‌های مولد تخاصمی(Generative Adversarial Networks)
٥. شبکه‌های تابع پایه شعاعی(Radial Basis Function Networks)
٦. پرسپترون‌های چندلایه(Multilayer Perceptrons)
۷. نگاشت‌های خودسازمان‌دهنده(Self-Organizing Maps)
۸. شبکه‌های باور عمیق(Deep Belief Networks)
۹. ماشین‌های بولتزمن محدود شده(Restricted Boltzmann Machines)
۱۰. اتوانکدرها یا خودرمزگذارها(AutoEncoders)

الگوریتم‌های یادگیری عمیق، تقریباً با هر نوع داده‌ای کار می‌کنند و برای حل مسائل پیچیده، به حجم زیادی از اطلاعات و قدرت محاسباتی بالا نیاز دارند. اکنون، اجازه دهید در عمق این ۱۰ الگوریتم برتر یادگیری عمیق غوطه‌ور شویم.

1- شبکه‌های عصبی کانولوشنی (CNN):

Convolutional Neural Networks یا به اختصار CNNها که با عنوان ConvNets نیز شناخته می‌شوند، از چندین لایه تشکیل شده‌اند و عمدتا برای پردازش تصویر و تشخیص اشیا استفاده می‌شوند. یان لکان(Yann LeCun) ، اولین CNN را در سال ۱۹۸۸ ایجاد کرد که LeNet نام داشت. از این شبکه برای تشخیص کاراکترهایی مانند کد پستی و ارقام استفاده می‌شد.

شبکه‌های عصبی کانولوشنی، به طور گسترده برای شناسایی تصاویر ماهواره‌ای، پردازش تصاویر پزشکی، پیش بینی سری‌های زمانی و تشخیص ناهنجاری‌ها مورد استفاده قرار می‌گیرند.

شبکه‌های عصبی کانولوشنی چگونه کار می‌کنند؟

CNNها چندین لایه دارند که ویژگی‌ها را پردازش و از داده‌ها استخراج می‌کنند.

لایه کانولوشن:

- این شبکه دارای یک لایه کانولوشن است که چندین فیلتر برای انجام عملیات کانولوشن دارد.

واحد خطی یکسوسازی شده:(Rectified Linear Unit)

- این شبکه‌ها دارای یک لایه ReLU برای انجام عملیات روی عناصر هستند. خروجی آن، یک نگاشت ویژگی یکسوسازی شده است.

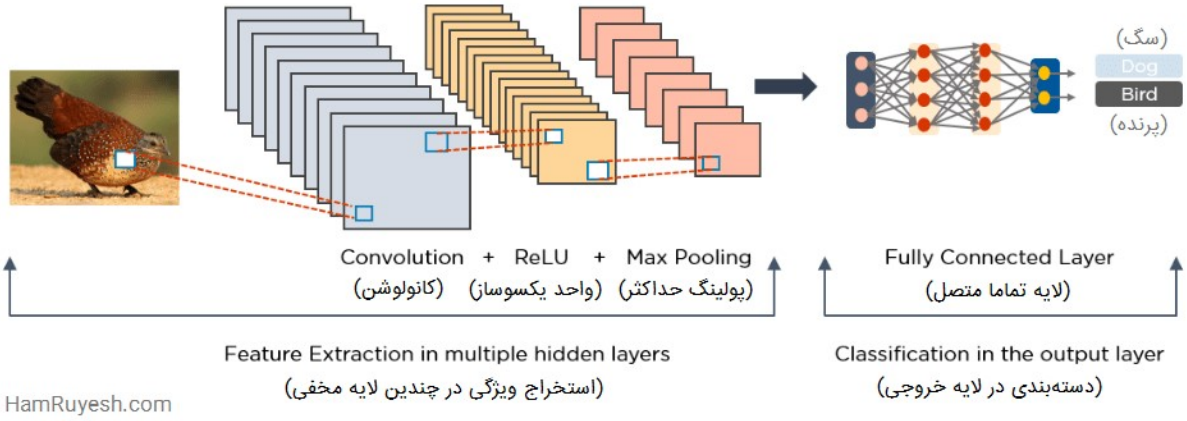
لایه پولینگ یا جمع آوری کننده:(pooling)

- نگاشت ویژگی یکسوسازی شده، در ادامه به یک لایه پولینگ داده می‌شود. پولینگ، یک عملیات نمونه برداری کاهشی است که ابعاد نگاشت ویژگی را کاهش می‌دهد.
- این لایه، سپس آرایه‌های دو بعدی حاصل از پولینگ نگاشت ویژگی را مسطح یا هموار (flatten) کرده و به یک بردار واحد، طولانی، پیوسته و خطی تبدیل می‌کند.

لایه تماما متصل:(fully connected)

- یک لایه تماما متصل، هنگامی ایجاد می‌شود که ماتریس هموار شده از لایه پولینگ، به عنوان ورودی داده شود، که تصاویر را دسته بندی و شناسایی می‌کند.

شکل زیر، نمونه‌ای از پردازش تصویر توسط شبکه عصبی کانولوشنی را نشان می‌دهد.



2-شبکه‌های حافظه طولانی کوتاه مدت:(LSTM)

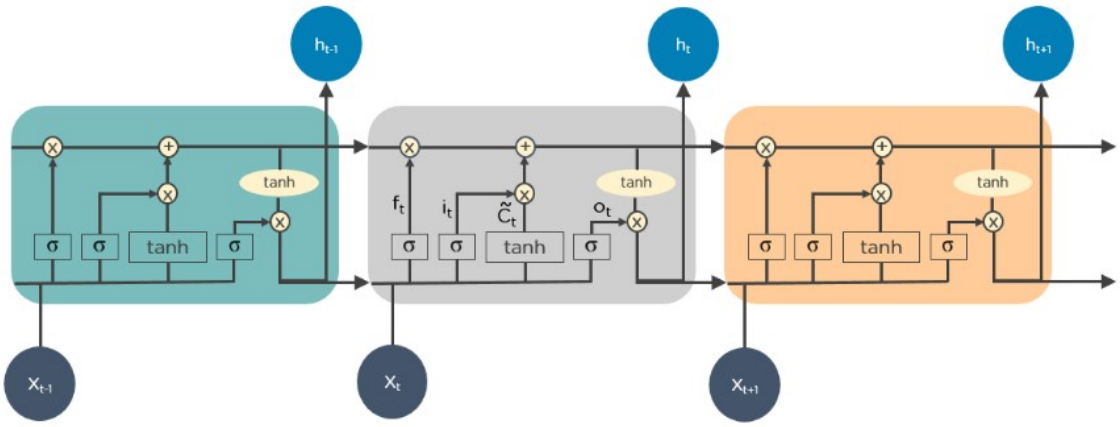
شبکه های Long Short Term Memory یا به اختصارLSTM ها، نوعی شبکه عصبی بازگشتی (RNN) هستند که می‌توانند وابستگی‌های طولانی مدت را یاد بگیرند و به خاطر بسپارند. به یاد آوردن اطلاعات گذشته برای دوره‌های متناوب طولانی، رفتار پیش فرض آن هاست.

این شبکه‌ها اطلاعات را در طول زمان حفظ می‌کنند و در پیش بینی سری‌های زمانی مفید هستند زیرا ورودی‌های قبلی را به خاطر می‌آورندLSTM .ها دارای یک ساختار زنجیره‌ای هستند که در آن، چهار لایه در تعامل با هم، به روشی منحصر به فرد ارتباط برقرار می‌کنند. علاوه بر پیش بینی‌های سری زمانی، LSTMها معمولاً برای تشخیص گفتار، ترکیب موسیقی و توسعه داروسازی استفاده می‌شوند.

LSTMها چگونه کار می‌کنند؟

- ابتدا، قسمت‌های نامرتبط حالت قبلی را فراموش می‌کنند
- سپس، به صورت انتخابی، مقادیر حالت سلول (cell-state) را به روز می‌کنند
- در نهایت، قسمت‌های مشخصی از حالت سلول را به عنوان خروجی تحویل می‌دهند.

شکل زیر، نحوه عملکردLSTM ها را نشان می‌دهد:

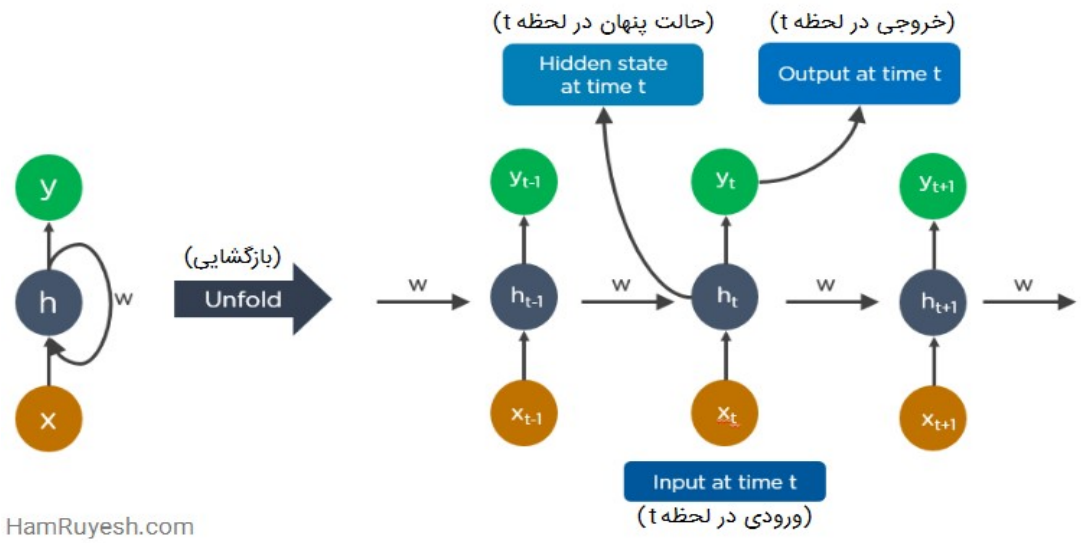


3-شبکه‌های عصبی بازگشتی:(RNN)

Recurrent Neural Networksیا به اختصار RNNها، دارای اتصالاتی هستند که حلقه‌هایی جهت‌دار را تشکیل می‌دهند، که به خروجی LSTM اجازه می‌دهد تا به عنوان ورودی به فاز فعلی تغذیه شود.

خروجیLSTM ، ورودی مرحله فعلی می‌شود و می‌تواند ورودی‌های قبلی را با توجه به حافظه داخلی خود حفظ کندRNN .ها معمولاً برای زیرنویس تصاویر، تجزیه و تحلیل سری‌های زمانی، پردازش زبان طبیعی، تشخیص دست خط و ترجمه ماشین استفاده می‌شوند.

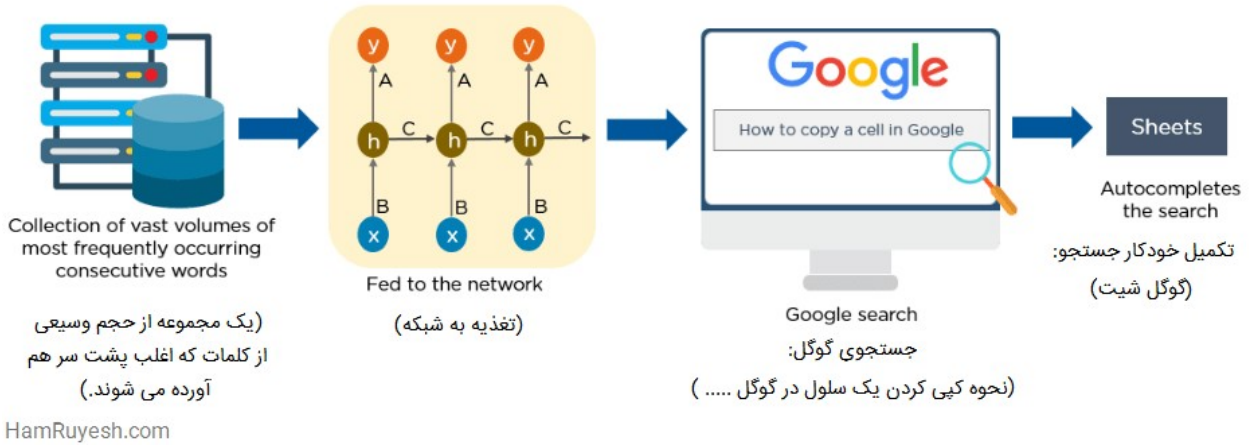
یک RNN باز شده به این شکل است:



شبکه‌های عصبی بازگشتی چگونه کار می‌کنند؟

- خروجی در لحظه t-1 به ورودی در لحظه t تغذیه می‌شود.
- به طور مشابه، خروجی در لحظه t به ورودی در لحظه t+1 تغذیه می‌شود.
- RNNها می‌توانند ورودی‌هایی با هر طول را پردازش کنند.
- محاسبات، برای اطلاعات تاریخی انجام می‌شود و اندازه مدل با اندازه ورودی افزایش نمی‌یابد.

شکل زیر، مثالی از نحوه عملکرد ویژگی تکمیل خودکار توسط گوگل را نشان می‌دهد:



4-شبکه‌های مولد تخاصمی:(GAN)

Generative Adversarial Networksیا به اختصارGAN ها، الگوریتم‌های یادگیری عمیق مولد هستند که نمونه‌های جدیدی از داده‌ها را شبیه به داده‌های آموزشی ایجاد می‌کنند .GAN دو بخش اصلی دارد:

یک مولد (generator) که یاد می‌گیرد داده‌های جعلی تولید کند و یک متمایزگر یا تشخیص دهنده(discriminator) ، که از این اطلاعات نادرست درس می‌گیرد.

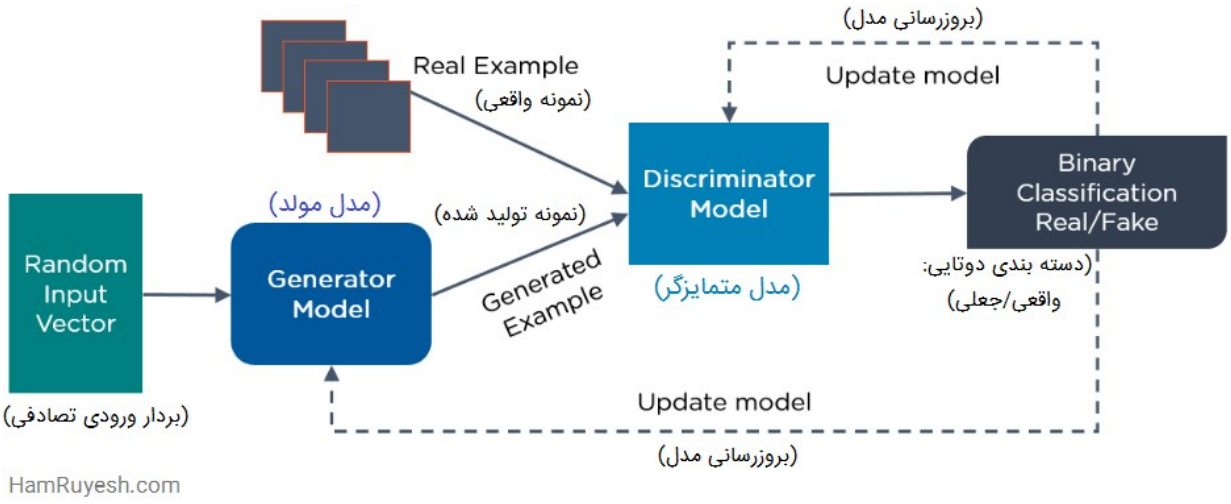
استفاده ازGAN ها طی یک دوره زمانی بسیار افزایش یافته است. از این شبکه‌ها می‌توان برای بهبود تصاویر اخترشناسی و شبیه سازی عدسی گرانشی برای تحقیقات در مورد ماده تاریک (dark-matter) استفاده کرد. توسعه دهندگان بازی‌های ویدئویی، از GANبرای ارتقای بافت‌های با وضوح پایین و دو بعدی در بازی‌های ویدئویی قدیمی با بازآفرینی آن‌ها در وضوح ۴ Kیا بالاتر، از طریق آموزش تصاویر استفاده می‌کنند.

GANها به تولید تصاویر واقعی و شخصیت‌های کارتونی، ایجاد عکس از چهره انسان و ارائه اشیاء سه بعدی کمک می‌کنند.

GAN چگونه کار می‌کند؟

- بخش متمایزگر، یاد می‌گیرد که بین داده‌های جعلی مولد و داده‌های نمونه واقعی، تمایز قائل شود.
- در طول آموزش اولیه، مولد، داده‌های جعلی تولید می‌کند، و متمایزگر به سرعت یاد می‌گیرد که بگوید این اطلاعات غلط است.
- GANنتایج را به مولد و متمایزگر ارسال می‌کند تا مدل را به روز کند.

در شکل زیر، نحوه عملکردGAN ها نشان داده شده است:



5-شبکه‌های تابع پایه شعاعی:(RBFN)

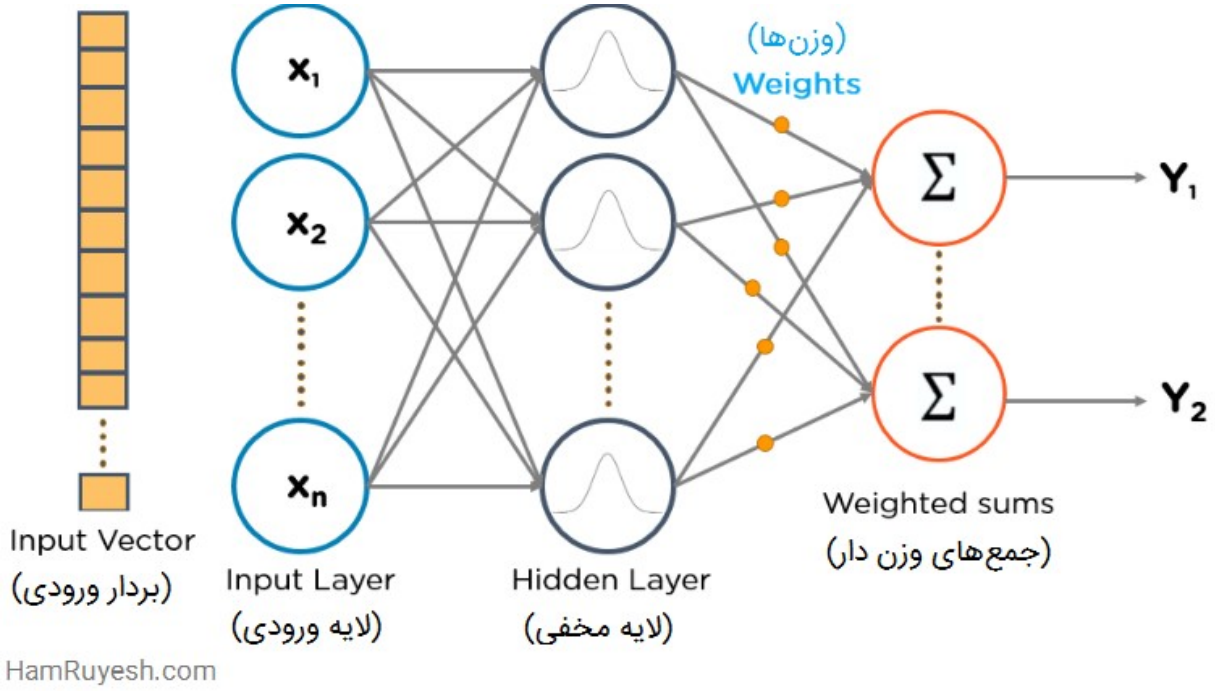
Radial Basis Function Networksیا به اختصارRBFN ها، انواع خاصی از شبکه‌های عصبی پیشخور هستند که از توابع پایه شعاعی به عنوان توابع فعال سازی استفاده می‌کنند. این شبکه‌ها یک لایه ورودی، یک لایه مخفی و یک لایه خروجی دارند و بیشتر برای دسته بندی، رگرسیون و پیش بینی سری‌های زمانی استفاده می‌شوند.

RBFNها چگونه کار می‌کنند؟

- RBFNها با اندازه گیری شباهت ورودی به نمونه‌های مجموعه آموزشی، کار دسته بندی را انجام می‌دهند.
- RBFNها یک بردار ورودی دارند که به لایه ورودی داده می‌شود. این شبکه‌ها همچنین دارای لایه‌ای از نورون‌های تابع پایه شعاعی (RBF) هستند.
- تابع، مجموع وزن دار ورودی‌ها را پیدا می‌کند و در لایه خروجی، به ازای هر دسته یا کلاس، یک گره وجود دارد.

- نورون‌های موجود در لایه پنهان، شامل توابع انتقال گوسی هستند که خروجی‌های آن‌ها، با فاصله از مرکز نورون، نسبت معکوس دارد.
- خروجی شبکه، یک ترکیب خطی از توابع پایه شعاعی ورودی و پارامترهای نورون است.

یک نمونه RBFN در شکل زیر به نمایش در آمده است:



6-پرسپترون‌های چندلایه:(MLP)

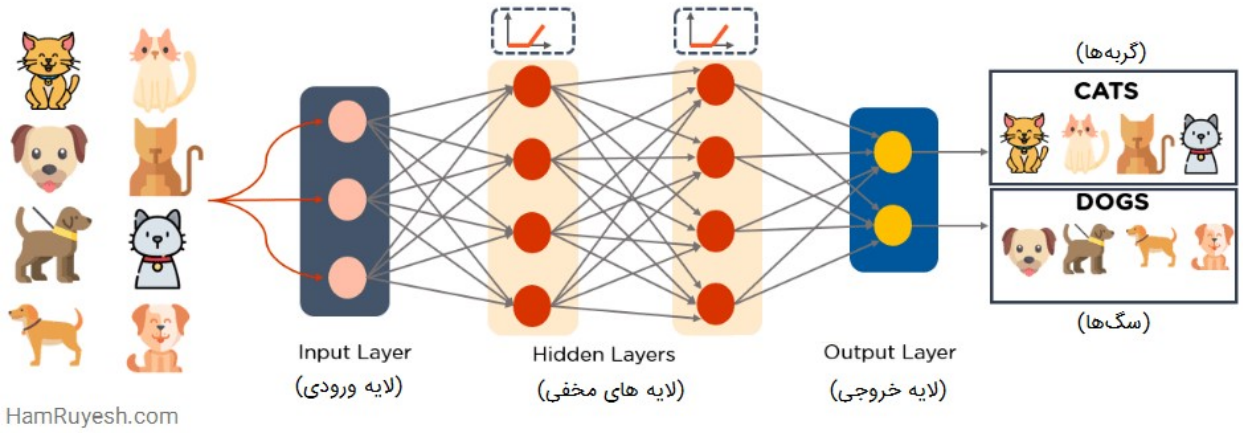
MultiLayer Perceptronsیا به اختصار MLPها، یک گزینه عالی برای شروع یاد گرفتن فناوری یادگیری عمیق به حساب می‌آیند.

MLPها متعلق به دسته شبکه‌های عصبی پیشخور با چندین لایه پرسپترون و شامل توابع فعال سازی هستندMLP .ها از یک لایه ورودی و یک لایه خروجی تشکیل شده‌اند که تمامی آن‌ها به هم متصل شده‌اند. این شبکه‌ها دارای تعداد لایه ورودی و خروجی یکسانی هستند اما ممکن است چندین لایه مخفی داشته باشند. این شبکه‌ها می‌توانند برای ساختن نرم افزارهای تشخیص گفتار، تشخیص تصاویر و ترجمه ماشین مورد استفاده قرار بگیرند.

MLPها چگونه کار می‌کنند؟

- MLPها، داده‌ها را به لایه ورودی شبکه تغذیه می‌کنند. لایه‌های نورون‌ها در یک گراف به هم متصل می‌شوند تا سیگنال در یک جهت عبور کند.
- MLPها ورودی را با وزنی که بین لایه ورودی و لایه‌های مخفی وجود دارد محاسبه می‌کنند.
- MLPها با استفاده از توابع فعال سازی تعیین می‌کنند که کدام گره‌ها شلیک کنند. توابع فعال سازی شامل ReLU ، تابع سیگموئید (sigmoid) و tanh می‌باشند.
- MLPها مدل را برای درک همبستگی و یادگیری وابستگی‌های بین متغیرهای مستقل و هدف در یک مجموعه داده آموزشی تعلیم می‌دهند.

در شکل زیر، نمونه‌ای از MLP نشان داده شده است. این شبکه، وزن‌ها و بایاس را محاسبه می‌کند و توابع فعال سازی مناسب را برای دسته بندی تصاویر گربه‌ها و سگ‌ها اعمال می‌کند.



7-نگاشت‌های خودسازمان‌دهنده:(SOM)

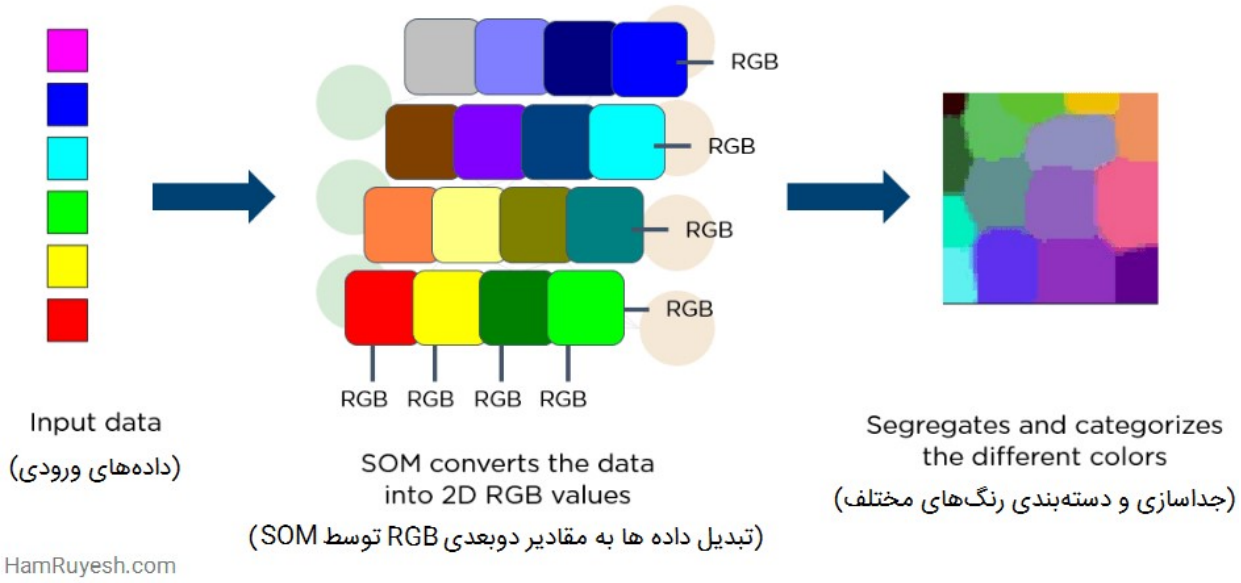
پروفسور تیوو کوهونن(Teuvo Kohonen) ، Self-Organizing Mapsیا به اختصارSOM ها را اختراع کرد. این نگاشت‌ها به مصورسازی و تجسم داده‌ها (data visualization) کمک می‌کند تا ابعاد داده‌ها را از طریق خود سازماندهی شبکه‌های عصبی مصنوعی کاهش دهند.

مصورسازی داده‌ها تلاش می‌کند مسئله‌ای را حل کند که انسان‌ها نمی‌توانند به راحتی داده‌های آن را با ابعاد بالا تجسم کنند . SOMها برای کمک به کاربران به منظور درک این اطلاعات با ابعاد بالا ایجاد شده‌اند.

SOMها چگونه کار می‌کنند؟

- SOMها وزن‌های هر گره را مقداردهی اولیه کرده و یک بردار را از داده‌های آموزشی به صورت تصادفی انتخاب می‌کنند.
- SOMها هر گره را بررسی می‌کنند تا دریابند کدام وزن‌ها محتمل ترین بردار ورودی هستند. گره برنده، بهترین واحد تطابق Best Matching Unit) یا (BMU نامیده می‌شود.
- SOMها همسایگی BMU را کشف می‌کنند و تعداد همسایه‌ها با گذشت زمان کاهش می‌یابد.
- SOMها یک وزن برنده را به بردار نمونه تحویل می‌دهند. هرچه یک گره به BMU نزدیک تر باشد، وزن آن بیشتر تغییر می‌کند.
- هرچه همسایه از BMU دورتر باشد، کمتر یاد می‌گیردSOM .ها مرحله دوم را برای N تکرار انجام می‌دهند.

در شکل زیر، تصویری از بردار ورودی با رنگ‌های مختلف را مشاهده می‌کنید. این داده‌ها به یک SOM تغذیه می‌شوند، که سپس داده‌ها را به مقادیر دوبعدی RGB تبدیل می‌کند. در نهایت، رنگ‌های مختلف را جدا و دسته بندی می‌کند.



8-شبکه‌های باور عمیق:(DBN)

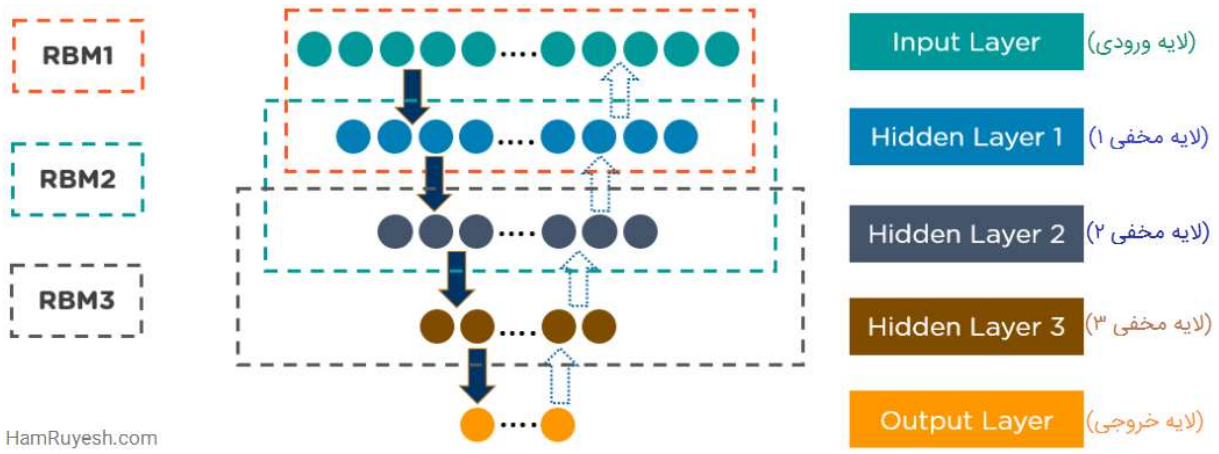
Deep Belief Networksیا به اختصارDBN ها، مدل‌های مولدی هستند که از لایه‌های متعددی از متغیرهای تصادفی و نهفته تشکیل شده‌اند. متغیرهای نهفته (latent variables) دارای مقادیر دوتایی هستند و اغلب به آن‌ها واحدهای مخفی گفته می‌شود.

DBNها، مجموعه‌ای از ماشین‌های بولتزمن با اتصالاتی بین لایه‌ها هستند و هر لایهRBM ، هم با لایه قبلی و هم با لایه بعدی ارتباط برقرار می‌کند. شبکه‌های باور عمیق (DBNs) برای تشخیص تصویر، تشخیص ویدئو و ضبط حرکت داده‌ها یا همان موشن کپچر (motion capture) استفاده می‌شود.

DBNها چگونه کار می‌کنند؟

- DBNها توسط الگوریتم‌های یادگیری حریصانه (greedy) آموزش می‌بینند. الگوریتم یادگیری حریصانه از یک رویکرد لایه به لایه برای یادگیری وزن‌های مولد بالا به پایین استفاده می‌کند.
- DBNها مراحل نمونه برداری گیبز (Gibbs sampling) را در دو لایه مخفی بالایی اجرا می‌کنند. در این مرحله یک نمونه از RBM که توسط دو لایه مخفی بالایی تعریف شده است، تهیه می‌شود.
- DBNها با استفاده از یک گذر (pass) از نمونه برداری اجدادی (ancestral sampling) از بقیه مدل، نمونه‌ای از واحدهای قابل مشاهده تهیه می‌کنند.
- DBNها یاد می‌گیرند که مقادیر متغیرهای نهفته در هر لایه را می‌توان با یک گذر از پایین به بالا استنباط کرد.

شکل زیر، نمونه‌ای از معماری DBN را نشان می‌دهد:



9-ماشین‌های بولتزمن محدود شده:(RBM)

Restricted Boltzmann Machinesیا به اختصارRBM ها، که توسط جفری هینتون (Geoffrey Hinton) توسعه یافته‌اند، شبکه‌های عصبی تصادفی هستند که می‌توانند از یک توزیع احتمال روی مجموعه‌ای از ورودی‌ها یاد بگیرند.

از این الگوریتم یادگیری عمیق، برای کاهش ابعاد، دسته بندی، رگرسیون، پالایش گروهی(collaborative filtering) ، یادگیری ویژگی‌ها و مدل سازی موضوع (topic modeling) استفاده می‌شودRBM .ها، بلوک‌های اصلیDBN ها را تشکیل می‌دهند.

RBMها از دو لایه تشکیل شده‌اند:

- واحدهای قابل مشاهده(visible)
- واحدهای مخفی(hidden)

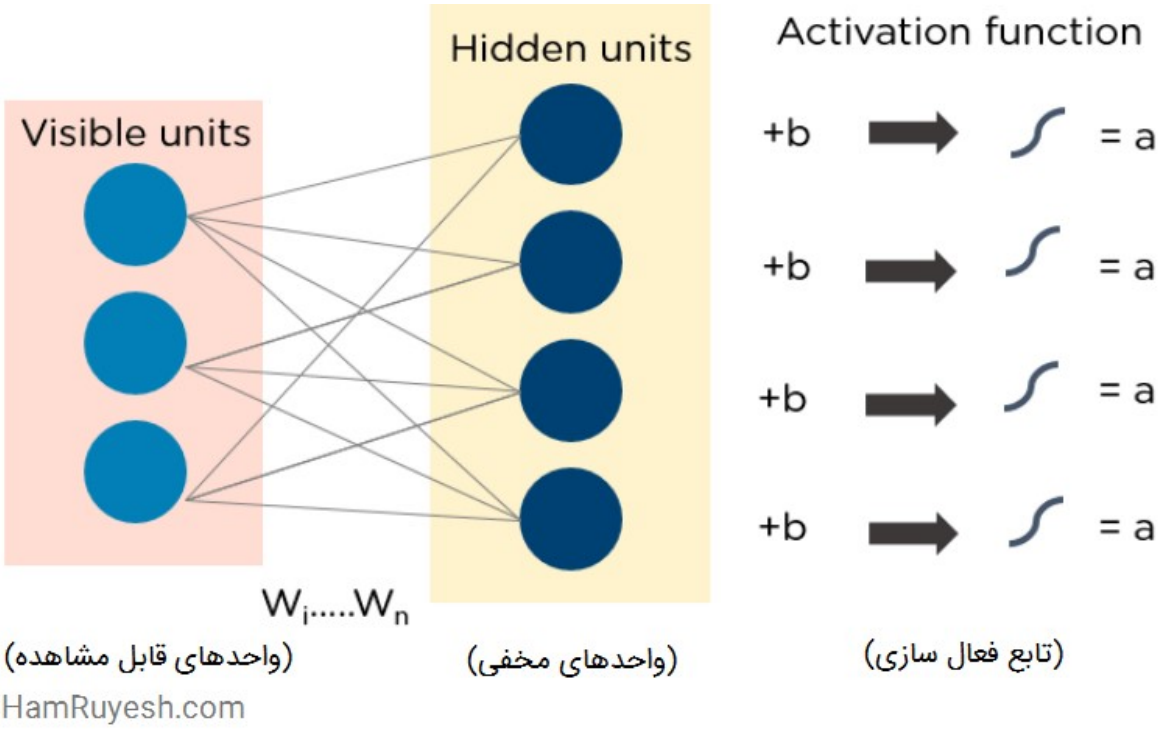
هر واحد قابل مشاهده، به تمام واحدهای مخفی متصل شده است RBM .ها دارای یک واحد بایاس هستند که به تمام واحدهای قابل مشاهده و واحدهای مخفی متصل است و هیچ گره خروجی ندارند.

RBMها چگونه کار می‌کنند؟

RBMها دارای دو فاز هستند: گذر رو به جلو و گذر رو به عقب.

- RBMها ورودی‌ها را می‌پذیرند و آن‌ها را به مجموعه‌ای از اعداد ترجمه می‌کنند که ورودی‌ها را در گذر رو به جلو رمزگذاری می‌کند.
- RBMها هر ورودی را با یک وزن منحصر به فرد و یک بایاس کلی ترکیب کرده و خروجی را برای بازسازی به لایه قابل مشاهده منتقل می‌کنند.
- در لایه قابل مشاهده، RBMبازسازی را با ورودی اصلی مقایسه می‌کند تا کیفیت نتیجه را تجزیه و تحلیل کند.

شکل زیر نحوه عملکردRBM ها را به تصویر کشیده است:



10-اتوانکدرها یا خودرمزگذارها:

خودرمزگذارها(Autoencoders) ، نوع خاصی از شبکه عصبی پیشخور هستند که در آن‌ها، ورودی و خروجی یکسان است. جفری هینتون در دهه ۱۹۸۰ خودرمزگذارها را برای حل مسائل یادگیری بدون نظارت طراحی کرد.

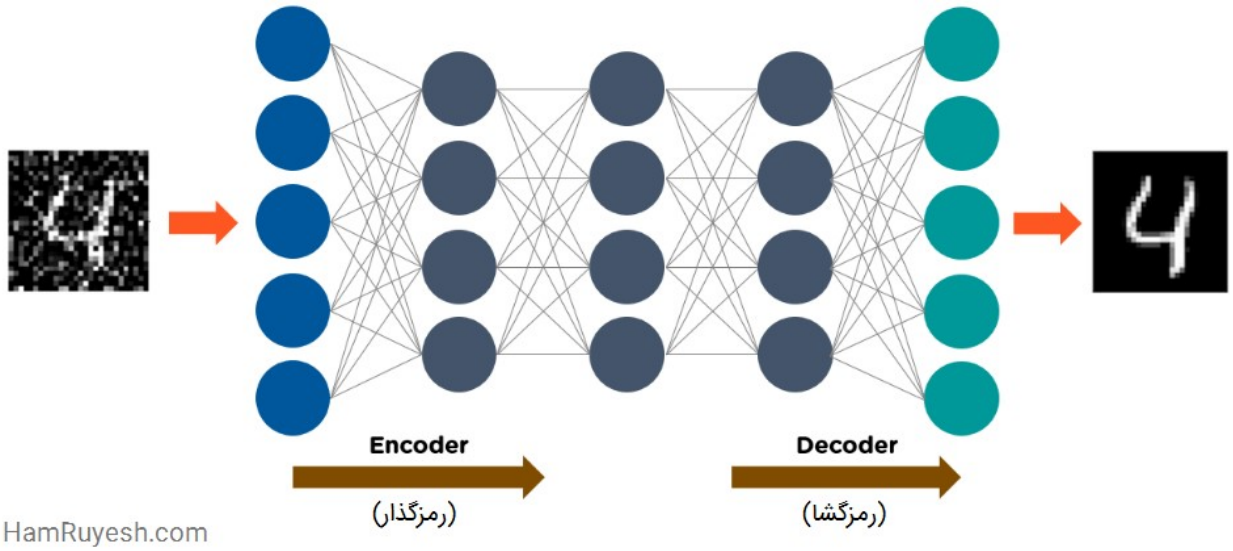
خودرمزگذارها شبکه‌های عصبی آموزش دیده‌ای هستند که داده‌ها را از لایه ورودی به خروجی همانندسازی (replicate) می‌کنند. از خودرمزگذارها برای اهدافی مانند اکتشافات دارویی، پیش بینی محبوبیت و پردازش تصویر استفاده می‌شود.

خودرمزگذارها چگونه کار می‌کنند؟

یک خودرمزگذار از سه قسمت اصلی تشکیل شده است: رمزگذار، رمز و رمزگشا.

- خودرمزگذارها به گونه‌ای ساختار یافته‌اند تا ورودی را دریافت کرده و آن را به یک بازنمایی متفاوت تبدیل کنند. سپس سعی می‌کنند ورودی اصلی را تا آنجا که ممکن است، به درستی بازسازی کنند.
- هنگامی که تصویری از یک رقم به وضوح قابل مشاهده نیست، آن تصویر به یک شبکه عصبی خودرمزگذار تغذیه می‌شود.
- خودرمزگذارها در ابتدا تصویر را رمزگذاری می‌کنند، سپس اندازه ورودی را به یک بازنمایی کوچک تر کاهش می‌دهند.
- در نهایت، خودرمزگذار، تصویر را رمزگشایی می‌کند تا تصویر بازسازی شده را تولید کند.

شکل زیر نحوه عملکرد خودرمزگذارها را نشان می‌دهد:



جمع بندی

یادگیری عمیق در پنج سال اخیر بسیار تکامل یافته است و الگوریتم‌های یادگیری عمیق در بسیاری از صنایع محبوبیت زیادی پیدا کرده‌اند. اگر به دنبال ورود به یک حرفه هیجان انگیز در زمینه علم داده هستید و می‌خواهید نحوه کار با الگوریتم‌های یادگیری عمیق را بیاموزید.

منبع: hamruyesh.com