

به نام خدا



دانشگاه آزاد اسلامی - واحد قزوین
دانشکده ی رایانه و فناوری اطلاعات

سمینار دوره کارشناسی ارشد هوش مصنوعی

موضوع:

یادگیری عمیق

دانشجو:

حمید یوسفی

استاد راهنما:

جناب آقای دکتر رشیدی کنعان

استاد سمینار:

جناب آقای دکتر میبدی

فهرست مطالب

- * اهداف گزارش
- * تاریخچه
- * معرفی یادگیری عمیق
- * انگیزه های معماری عمیق
- * مزایای معماری عمیق
- * شبکه عصبی عمیق
- * انکدر خودکار
- * یادگیری نظارت نشده یا نیمه نظارتی و الگوریتم های عمیق
- * آموزش در معماری عمیق برای تشخیص تصویر
- * شبکه باور عمیق
- * کاربردها
- * نتایج حاصل شده برای الگوریتم های با یادگیری عمیق
- * نتیجه گیری

اهداف گزارش

* معرفی نظریه های مربوط به معماری عمیق

* بیان ایده شبکه عصبی عمیق

* معرفی معماری جدید شبکه های عصبی عمیق کانولوشن

* بررسی الگوریتم های یادگیری متکی بر تجزیه محلی

* بررسی چند الگوریتم مبتنی بر گرادیان

* بیان یادگیری بازنمایی و نقش آن در آموزش عمیق

* بیان نقش ماشین بولزمن در شبکه باور عمیق

تاریخچه

* یادگیری عمیق به طور خاص برای شبکه های عصبی مصنوعی ساخته شد و تاریخ استفاده از معماری های عمیق روی شبکه های عصبی به سال 1980، توسط Fukushima برمیگردد.

* سال 1980 شخصی به نام Yann LeCun و همکارانش با استفاده از شبکه های عصبی استاندارد پس انتشار که در سال 1974 رایج شده بود، موفق به ساخت یک شبکه ی عصبی عمیق برای تشخیص کد پستی روی پاک نامه شدند.

* سال 1991، Jurgen Schmidhuber و همراه دانشجوی خود به نام Sepp Hochreiter توانست این کندی را حل کند. آنها تاثیر گرادینان را در ساختار عمیق کشف کردند.

* آقای Juyang Weng به همراه همکارانش نظریه ای را بیان داشت که بر طبق نشان داد که مغز انسان یک مدل یکپارچه و سه بعدی از اشیای را که میبیند در ذهنش نگاه میدارد.

تاریخچه (ادامه)

* در سال 1992، Schmidhuber، یک ایده بسیار مشابه برای حالت کلی تر از سلسله مراتب عمیق بدون نظارت از شبکه های عصبی Recurrent، و همچنین مزایای این شبکه برای بالا بردن سرعت یادگیری نظارت شده را نشان داد.

* در تشخیص گفتار بهتر استفاده شد و بهتر از الگوریتم های کم عمق و مدل های مرسوم برای پردازش گفتار مدل های گاوسی مخلوط GMM و مدل مخفی مارکف HMM به نتیجه رسید.
تا سال ۲۰۰۶ که رنسانسی در معماری عمیق به وجود آمد.

* Hinton شبکه باور عمیق را در سال ۲۰۰۶ ارائه داد. Deep Belief Network

* Bengio انکدر خودکار پراکنده را ارائه داد. Deep sparse autoencoders

* تیمی از گوگل با رهبری Andrew Ng و Jeff Dean، در سال ۲۰۱۲ شبکه عصبی ای را طراحی کردند که یاد گرفت که مفاهیم سطح بالاتر را نیز تشخیص دهد. به عنوان مثال، تشخیص گربه ها تنها از روی ویدئوهای موجود در سایت یوتیوب

معرفی یادگیری عمیق

* Deep Learning یا Learning feature hierarchies زیرمجموعه‌ای از یادگیری ماشین به حساب می‌آید، بر اساس یادگیری چندین سطحی از بازنمایی‌های مختلف می‌باشد که دارای یک ساختار سلسله‌مراتبی از ویژگی‌ها یا مفاهیم است که در آن مفاهیم سطح بالا از روی ویژگی‌های سطح پایین تعریف می‌شوند و مفاهیم سطح پایین نیز می‌توانند به تعریف مفاهیم سطح بالاتر کمک کنند، می‌باشد ([Bengio \(2009\)](#))

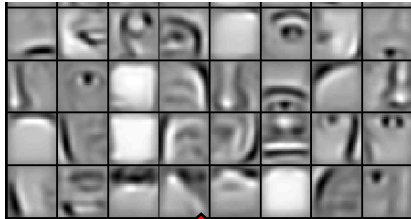
* در واقع به زبان ساده‌تر، هدف یادگیری عمیق، استخراج خود ویژگی‌ها به صورت هوشمند طی یک مرحله یادگیری است. تلاش‌ها در این زمینه (خصوصاً ساخت یک یادگیر با استفاده از شبکه‌های عصبی) تا قبل از سال ۲۰۰۶ موفقیت‌آمیز نبود. از این سال به بعد تلاش‌ها امیدوارکننده‌تر بود؛

معرفی یادگیری عمیق (ادامه)

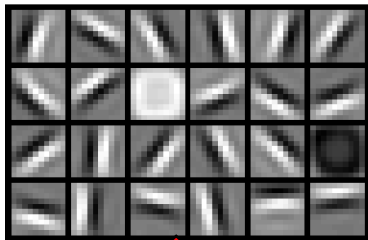
مدلی از اشیا موجود در تصویر



تقسیم بخشهای تصویر
(شی ها و لبه ها)



لبه تصاویر



پیکسل



معرفی یادگیری عمیق (ادامه)

❖ پرسش

❖ یادگیری عمیق دقیقا چیست؟

❖ چرا تعمیم بهتری نسبت به روشهای دیگر روی داده های مرتبط به تصویر و گفتار و برخی از انواع دیگر دارد؟

❖ پاسخ پرسش ها

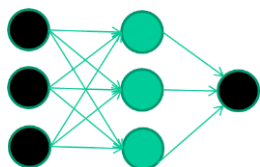
❖ یادگیری عمیق به معنی استفاده از یک شبکه عصبی با چندین لایه از نودها بین ورودی و خروجی ست

❖ لایه های بین ورودی و خروجی همانند مغز انسان ویژگی ها را طی یک سری از مراحل شناسایی و پردازش میکنند

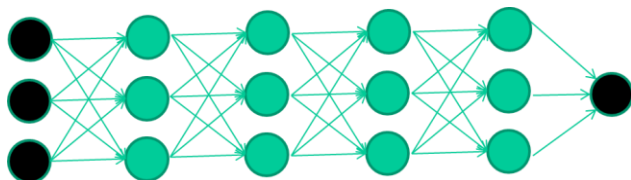
معرفی یادگیری عمیق (ادامه)

* شبکه های چند لایه بیش از ۲۵ سال است که وجود دارند، یادگیری عمیق چه نوآوری دارد؟

* همواره بهترین الگوریتم ها برای یادگیری وزن ها در شبکه های عصبی، معماری با ۱ لایه مخفی است.



* اما این الگوریتم ها برای آموزش وزن ها در شبکه های عصبی با چند لایه مخفی، به خوبی عمل نمیکنند.



* الگوریتم های یادگیری عمیق برای جبران این نقیصه ارایه شدند.

معرفی یادگیری عمیق (ادامه)

* یادگیری عمیق برای چه دسته از یادگیری ها مناسب است؟

* یادگیری با نظارت

* یادگیری نیمه نظارتی

* یادگیری خودآموخته

یادگیری با نظارت



خودرو



موتور سیکلت

این تصویر چیست؟



یادگیری نیمه نظارتی



تصویر بدون برچسب (فقط میگوییم تمام تصاویر مربوط به موتور سیکلت و خودرو است)



خودرو



موتور سیکلت

این تصویر چیست؟



یادگیری خود آموخته



تصاویر بدون برچسب (تصاویر تصادفی از اینترنت)



خودرو



موتور سیکلت

این تصویر چیست؟



انگیزه های معماری عمیق

* الهام بخشی از مغز

* پیچیدگی محاسباتی

* بهره وری آماری و به اشتراک گذاری قدرت آماری

* اشتراک گذاری قدرت آماری برای یادگیری چند وظیفه ای، یادگیری نیمه نظارتی، آموزش خود آموخته

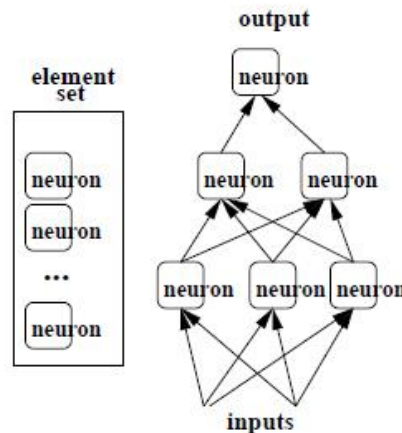
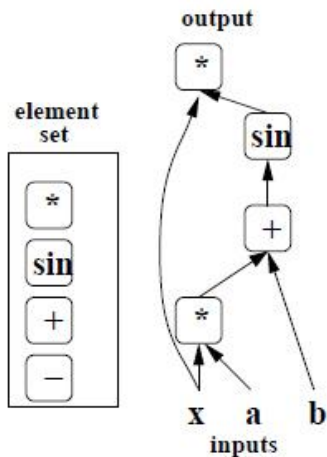
مزایای معماری عمیق

* تقسیم یک تابع به تعدادی عناصر محاسباتی

* بهبود عملکرد تعمیم به نسبت الگوریتم های با معماری کم عمق

* مسألی با منابع دانش اولیه کم

* بازنمایی فشرده در سطوح



معماری های عمیق

* شبکه های عصبی عمیق

* برای پردازش گفتار و مسأل با نظارت

* شبکه های پس انتشار Backpropagation

* شبکه های عصبی عمیق کانولوشن

○ ۱۹۸۰ Fukushima

○ ۱۹۹۸ Yann LeCun

○ ۲۰۰۶ Haffner و Bengio

○ برای مسأل طبقه بندی و تشخیص تصویر

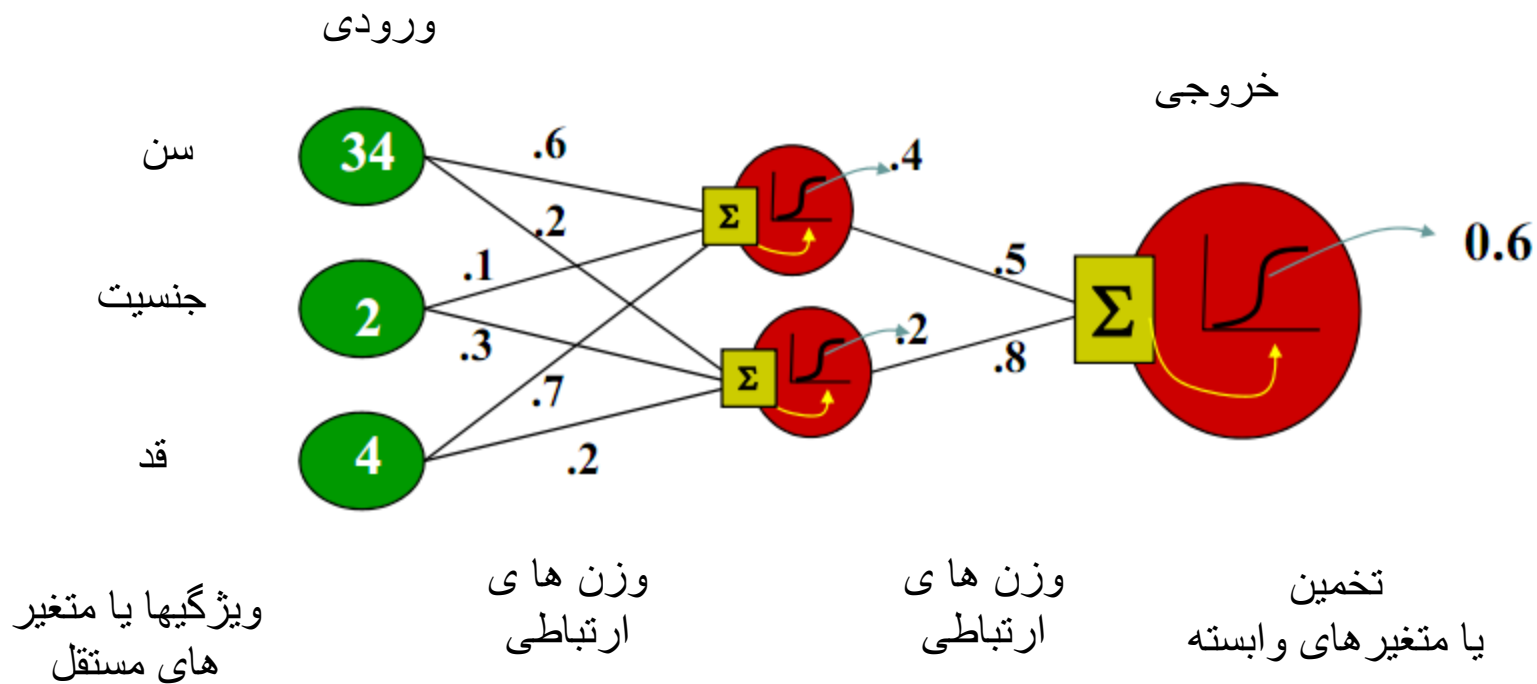
* شبکه های باور عمیق

* توسط Hinton در سال ۲۰۰۶

* برای یادگیری در مسأل بدون نظارت

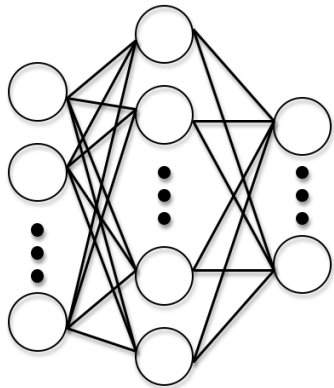
و غیره

شبکه های عصبی



شبکه عصبی عمیق

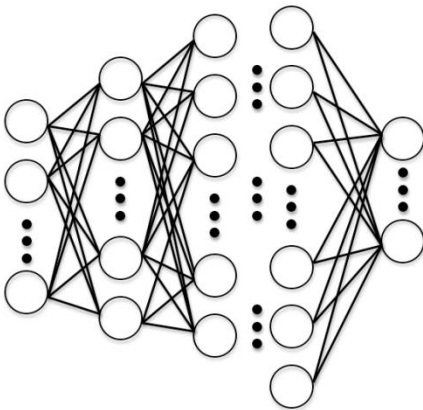
* می توان هر تابع در یک مجموعه فشرده یا compact را با شبکه های عصبی با یک لایه مخفی تخمین زد.



* برای شبکه عصبی عمیق

* میتوان تعداد لایه های عمیق را افزایش داد.

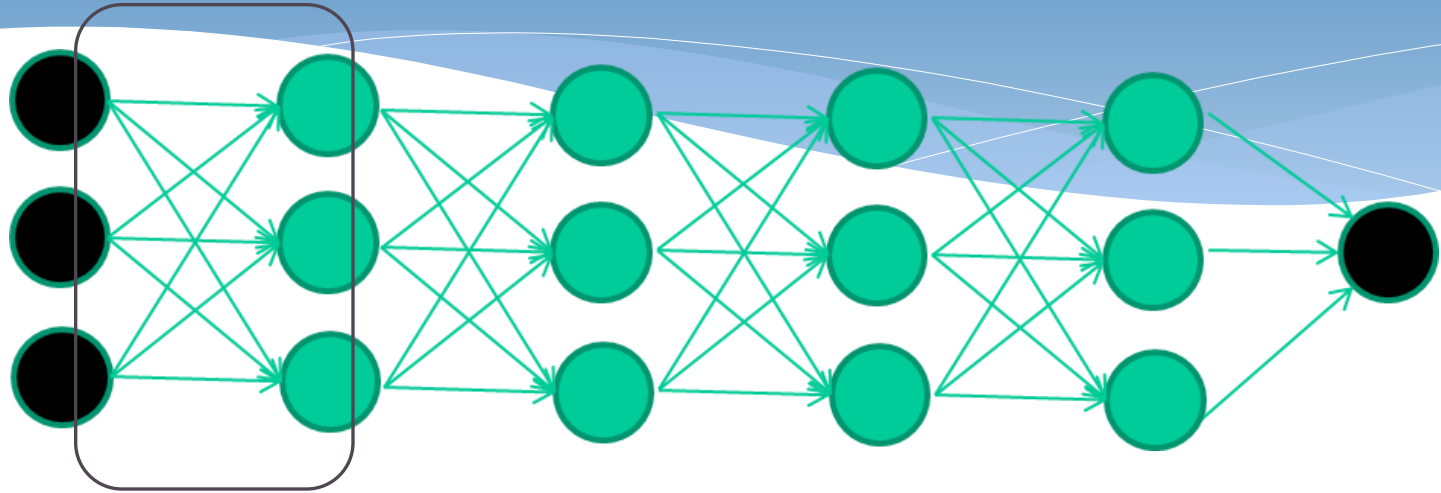
* هر لایه مخفی به صورت بدون نظارت آموزش می بیند.



* با همین ایده Hinton در سال ۲۰۰۶ شبکه عصبی

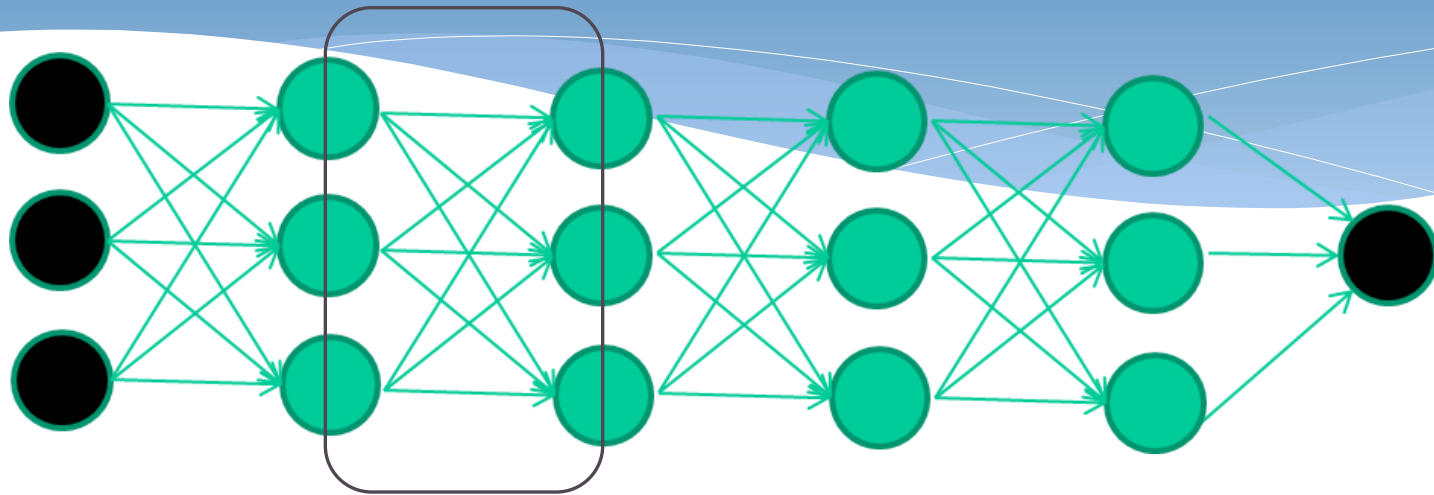
عمیق اش را ساخت.

آموزش عمیق برای شبکه عصبی عمیق با انکدر خودکار



آموزش لایه اول

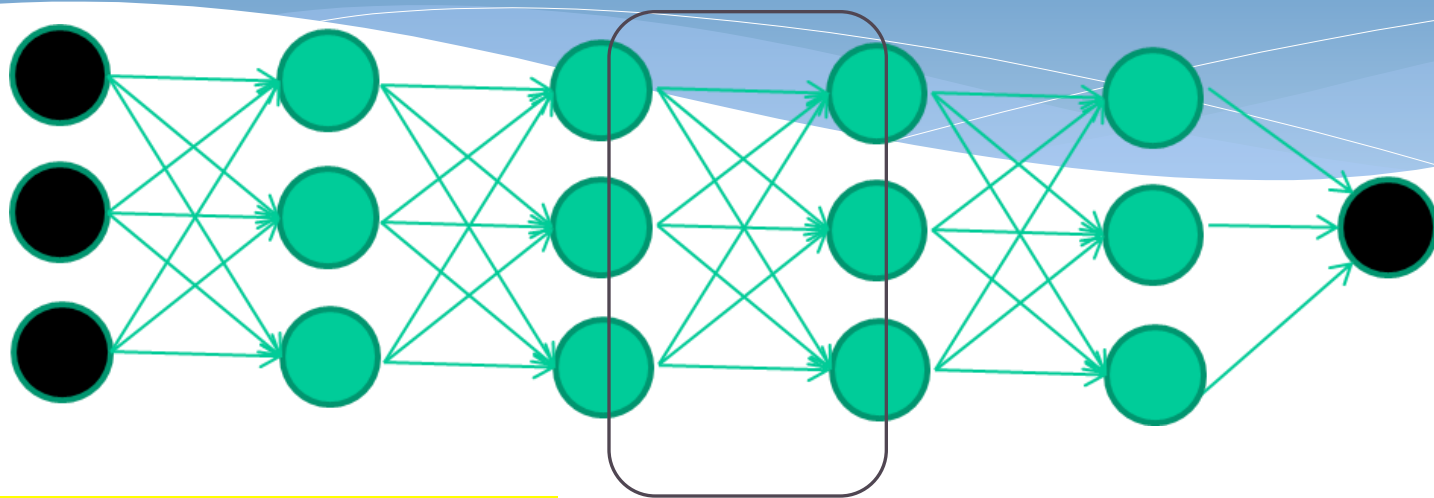
آموزش عمیق برای شبکه عصبی عمیق با انکدر خودکار



آموزش لایه اول

آموزش لایه دوم

آموزش عمیق برای شبکه عصبی عمیق با انکدر خودکار

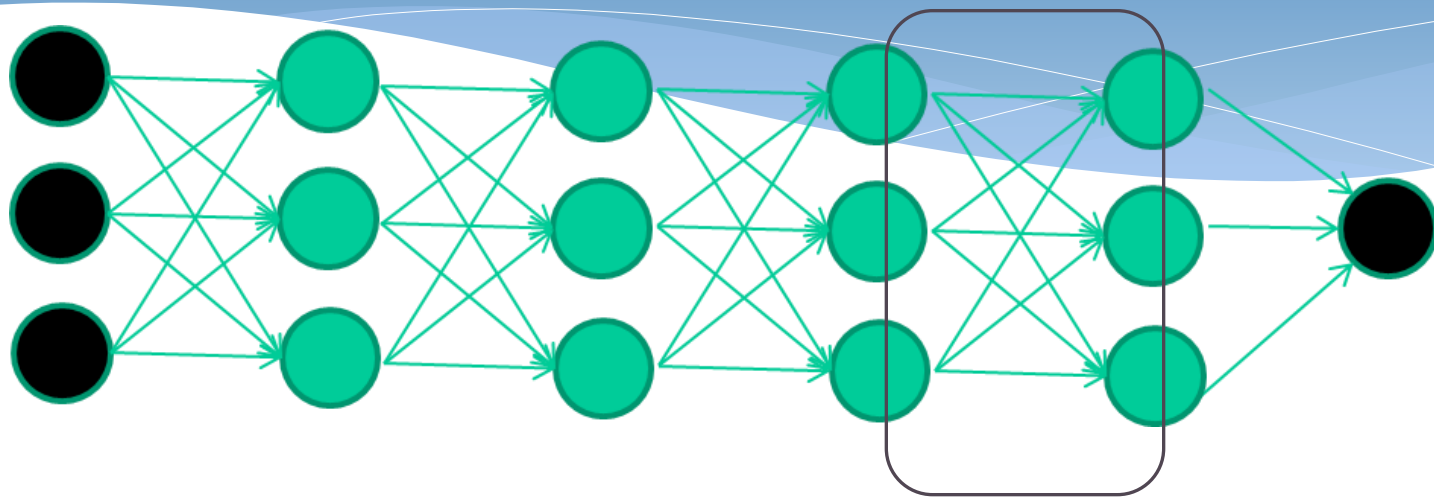


آموزش لایه اول

آموزش لایه دوم

آموزش لایه سوم

آموزش عمیق برای شبکه عصبی عمیق با انکدر خودکار



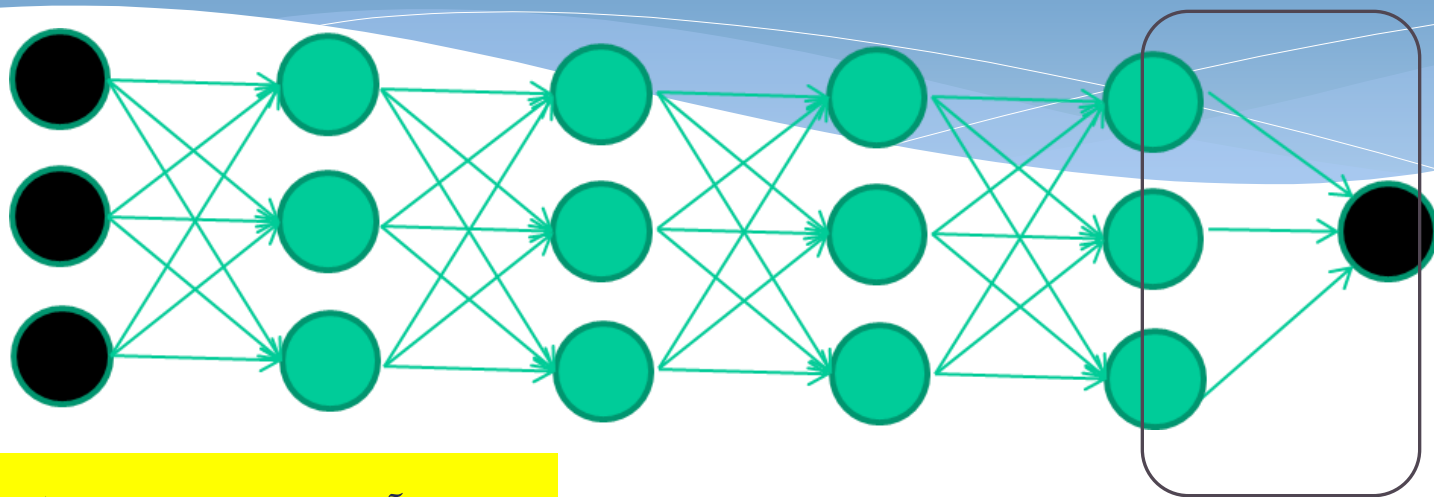
آموزش لایه اول

آموزش لایه دوم

آموزش لایه سوم

آموزش لایه چهارم

آموزش عمیق برای شبکه عصبی عمیق با انکدر خودکار



آموزش لایه اول

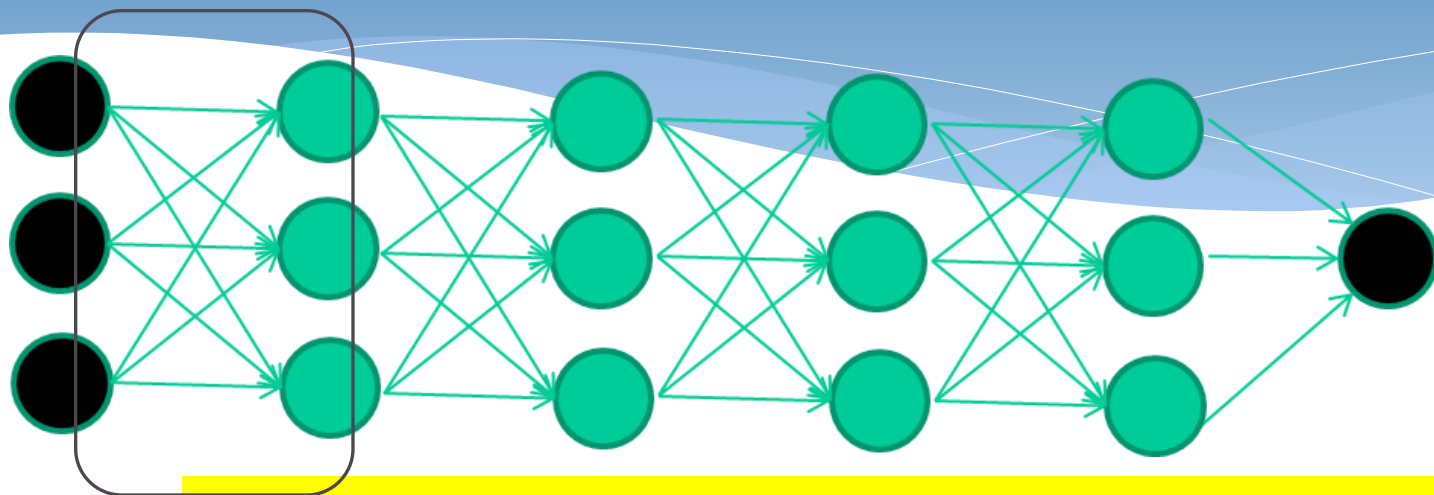
آموزش لایه دوم

آموزش لایه سوم

آموزش لایه چهارم

آموزش لایه نهایی

آموزش عمیق برای شبکه عصبی عمیق با انکدر خودکار

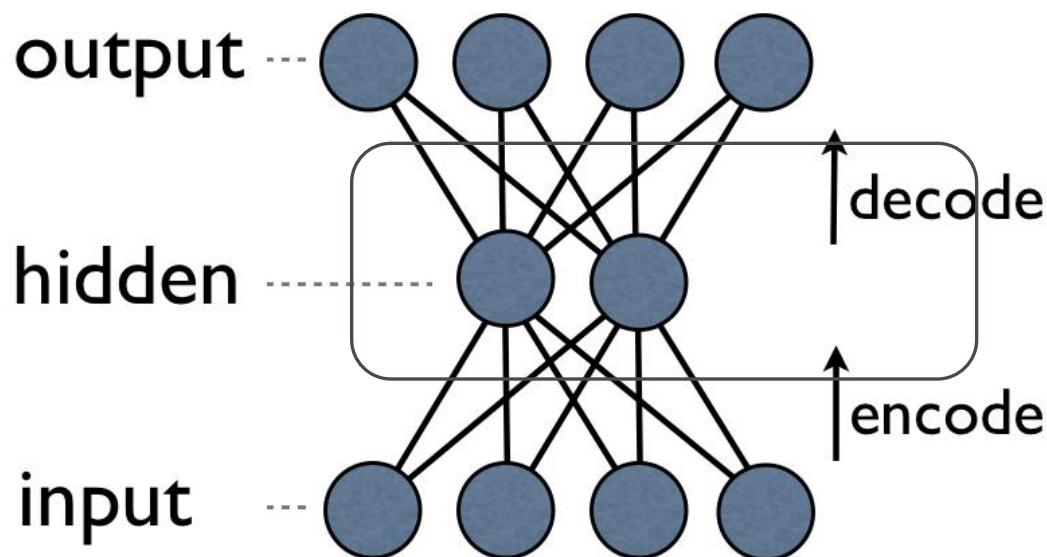


هر لایه به وسیله ی انکدر خودکار آموزش میبینند

در واقع ما شبکه را مجبور به یادگیری ویژگی های
خوب توصیف شده از لایه ی قبل کردیم

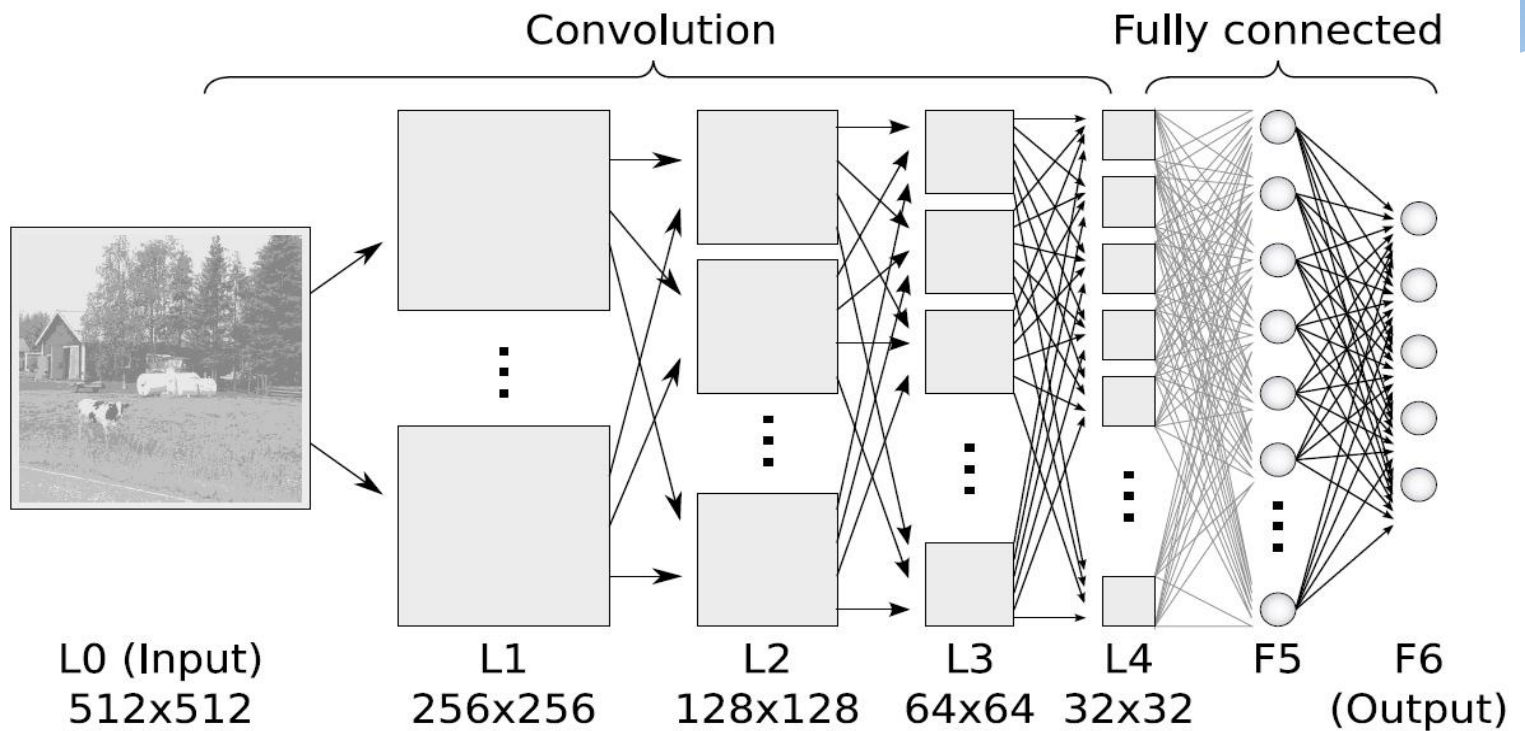
آموزش عمیق برای شبکه عصبی عمیق با انکدر خودکار

در آموزش انکدر های خودکار، با الگوریتم های تنظیم وزن به باز تولید ورودی ها میپردازند.



اتفاقی که در این معمار میافتد این است که به دلیل کم شدن ورودی ها ، لایه های مخفی مجبورند که ویژگی های خوب را بازنمایی و دیکد کنند.

معماری کانولوشن



یادگیری نظارت نشده یا نیمه نظارتی و الگوریتم های عمیق

* خطای آموزشی کمتر در هنگامی که مجموعه آموزشی بزرگ است.

- یک معیار آموزشی

- پیش آموزش لایه عاقلانه حریصانه

- مدل های گرافیکی بدون جهت و ماشین های بولتزمن

- ماشین بولتزمن محدود شده

- تخمین زننده واگرایی

- تخمین زننده حداکثر تصادفی

- واگرایی متقابل مداوم

* انجام پیش آموزش بدون نظارت روی داده های آموزشی

* به جای مقدار دهی اولیه در شبکه های عصبی از پیش آموزش استفاده شود

* لایه حریصانه-عاقلانه

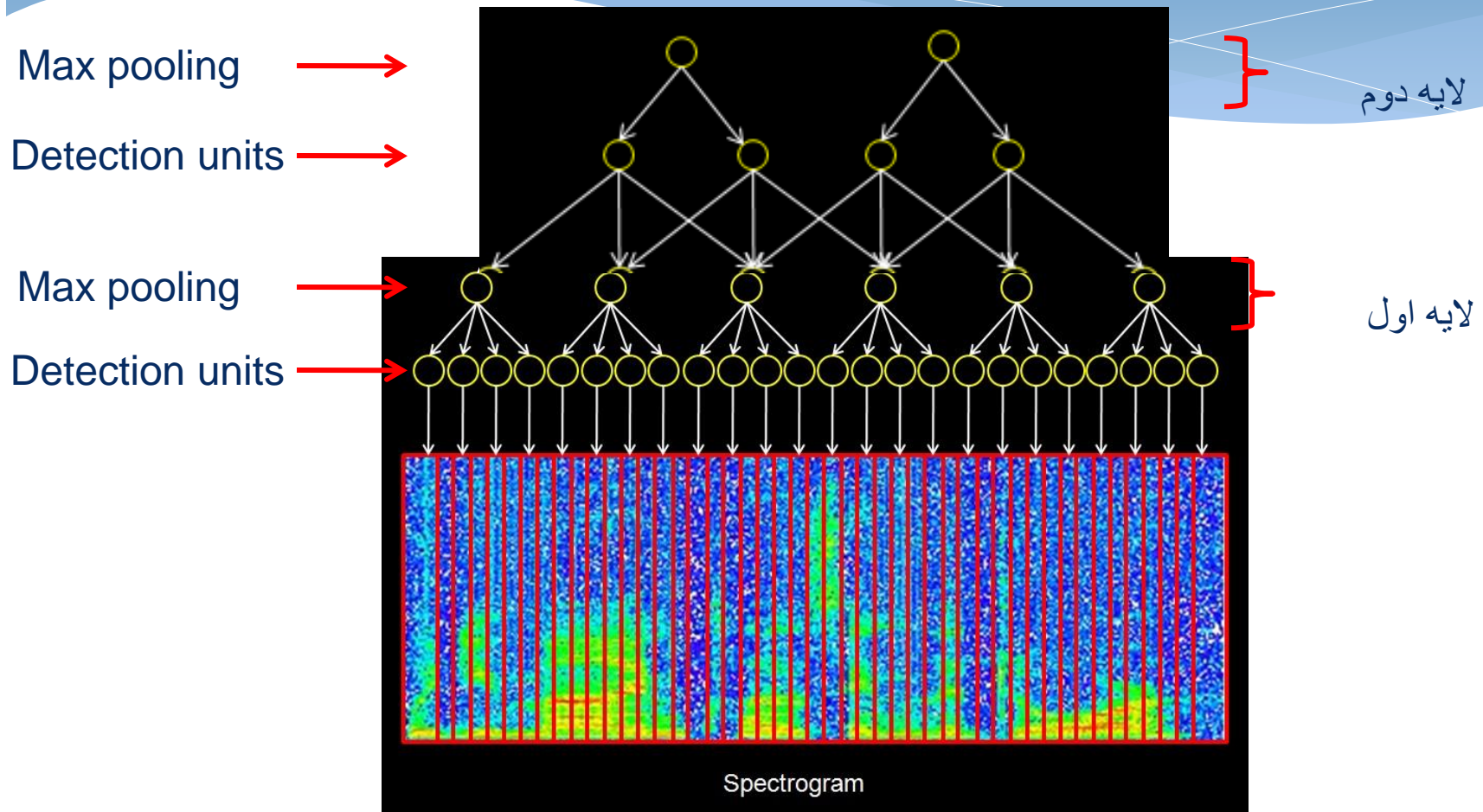
یادگیری نظارت نشده یا نیمه نظارتی و الگوریتم های عمیق (ادامه)

- * کدگذاری پراکنده انباشته
- * شبکه باور عمیق
- * انکدر های خودکار عمیق
- * Logistic regression
- * شبکه های عصبی
- * انکدر های خودکار پراکنده
- * انکدهای خودکار عمیق

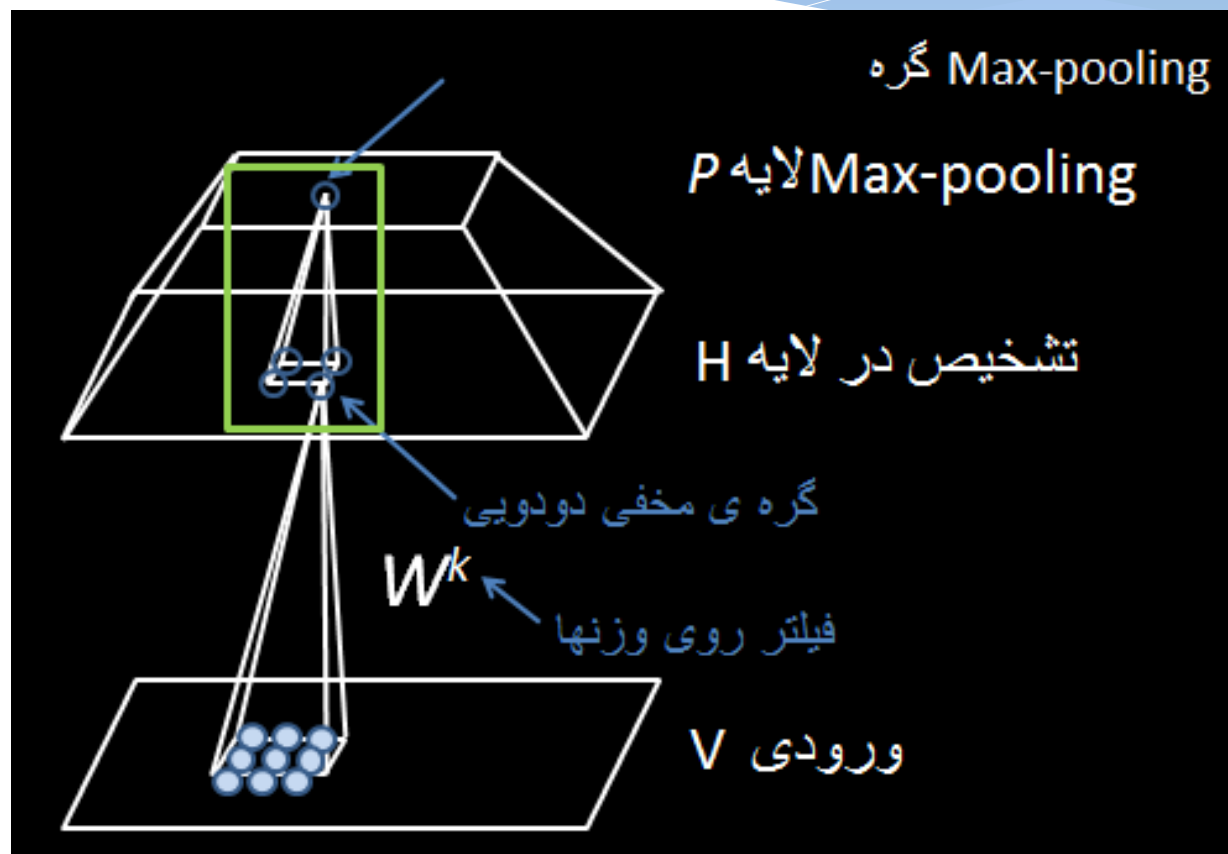
آموزش در معماری عمیق برای تشخیص تصویر

- * در سطوح پایین ترکیب ویژگیها مشترک با هم
- * در سطوح بالا بازنمایی انتزاعی تر
- * در سطوح پایین داده های خام ورودی
- * تغییر اندک اندک بازنمایی ها تا سطوح بالاتر
- * مراحل پردازش شامل شناسایی و تشخیص لبه ها و تشخیص اشکال ابتدایی موجود در تصویر است
- * به تدریج پردازش و تشخیص اشکال بصری پیچیده تر

شبکه ی باور عمیق کونولوشن برای پردازش صوت



شبکه ی باور عمیق کونولوشن برای پردازش تصویر



کاربردها

* پردازش گفتار و تشخیص گفتار

* پردازش زبانهای طبیعی

* تشخیص تصویر (فهم و درک اشیاء موجود در تصویر)، بینایی ماشین

➤ مسائل امنیتی

➤ تحلیل و تشخیص در عکس های پزشکی

➤ در راهنمایی و رانندگی و اتومبیل های خودکار

• برنامه های کاربردی معروف

➤ ترجمه گر آنلاین نرم افزار Skype شرکت مایکروسافت

➤ جستجوی صوت سایت Bing شرکت مایکروسافت

➤ نرم افزار های موجود در سیستم عامل اندروید و ios

خطای طبقه بندی بر روی مجموعه داده MNIST

Type	Preprocessing	Error rate
<u>Linear classifier</u>	Deskewing	7.6
<u>K-Nearest Neighbors</u>	Shiftable edges	0.52
<u>Support vector machine</u>	Deskewing	0.56
<u>Neural network</u>	None	0.35
<u>Convolutional neural network</u>	Width normalizations	0.23

نتایج آزمایشات دانشگاه استنفورد با یادگیری عمیق

Audio

TIMIT Phone classification	Accuracy
Prior art (Clarkson et al., 1999)	79.6%
Stanford Feature learning	80.3%

TIMIT Speaker identification	Accuracy
Prior art (Reynolds, 1995)	99.7%
Stanford Feature learning	100.0%

Images

CIFAR Object classification	Accuracy
Prior art (Krizhevsky, 2010)	78.9%
Stanford Feature learning	81.5%

NORB Object classification	Accuracy
Prior art (Ranzato et al., 2009)	94.4%
Stanford Feature learning	97.3%

Video

Hollywood2 Classification	Accuracy
Prior art (Laptev et al., 2004)	48%
Stanford Feature learning	53%
KTH	Accuracy
Prior art (Wang et al., 2010)	92.1%
Stanford Feature learning	93.9%

YouTube	Accuracy
Prior art (Liu et al., 2009)	71.2%
Stanford Feature learning	75.8%
UCF	Accuracy
Prior art (Wang et al., 2010)	85.6%
Stanford Feature learning	86.5%

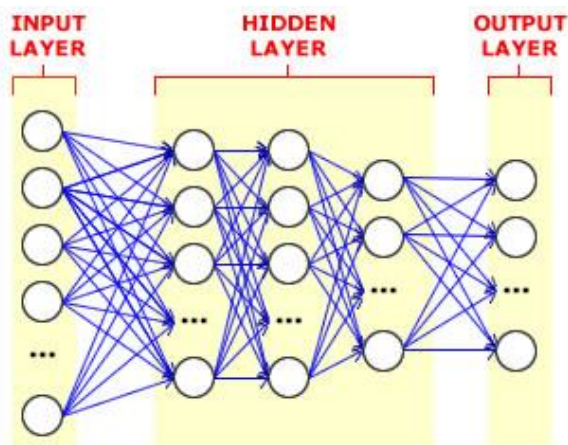
Multimodal (audio/video)

AVLetters Lip reading	Accuracy
Prior art (Zhao et al., 2009)	58.9%
Stanford Feature learning	65.8%

Other unsupervised feature learning records:
Pedestrian detection (Yann LeCun)
Different phone recognition task (Geoff Hinton)
PASCAL VOC object classification (Kai Yu)

و این که

- * یادگیری عمیق ، ایده ی به نسبت تازه ای محسوب میشود.
- * در این ارایه تنها اطلاعات اولیه درباره ی یادگیری عمیق گفته شود.
- * الگوریتم های یادگیری عمیق بسیاری وجود دارد.
- * انواع انکدر های خودکار، تغییرات روی معماری عمیق و الگوریتم های آموزشی جدید هر روزه معرفی میشود.



نتیجه گیری

- * برای مسال خودآموخته و نیمه نظارتی که داده های برچسب دار یا وجود ندارد و یا به اندازه ی کافی وجود ندارد این یادگیری مناسب است.
- * زمانیکه تعداد داده های آموزشی زیاد است استفاده از این رویکرد توصیه میشود.
- * یک معماری محاسباتی که به اندازه ی کافی عمق ندارد به تعداد زیادی عناصر محاسباتی نیاز دارد.
- * اگر الگوریتم های یادگیری تنها برتجزیه ی محلی متکی باشند، بعید است که به یک تعمیم خوب برای توابعی که دارای تعداد زیادی متغیر هست دست یابند.
- * شبکه های عصبی عمیق برای مسایل طبقه بندی مانند تشخیص شی در تصویر نتایج امیدوار کننده ای بدست آوردند.
- * افزایش قدرت تعمیم محلی با تلاش برای یادگیری توابع متفاوت در این الگوریتم ها بدست میآید.

پیشنهادهای آتی

* بررسی دلیل عدم موفقیت روشهای گرادیان استاندارد برای مسائل طبقه بندی مانند تشخیص

* استفاده از معماری عمیق در مسائل پزشکی (تشخیص نوع توده سرطانی و یا طبقه بندی یادگیری عمیق برضایعات کانونی کبد در سونوگرافی)

* پردازش های خودکار روی صوت (تشخیص گوینده در یک اتاق کنفرانس) Speaker

Diarization of Meetings Recordings

* استفاده از روش یادگیری عمیق برای پیشنهاد دوستی و ارتباط در شبکه های اجتماعی

* تشخیص رفتار های افراد توسط شبکه های کانولوشن روی تصاویر ۳ بعدی

منابع اصلی

- [19] Deep Learning of Representations, Yoshua Bengio and Aaron Courville, Dept. IRO, Université de Montréal C.P. 6128, Montreal, Qc, H3C 3J7, Canada ,2013
- [29] Learning Deep Architectures for AI, Yoshua Bengio , Dept. IRO, Université de Montréal C.P. 6128, Montreal, Qc, H3C 3J7, Canada ,2013
- [28] Lee, T.-S., & Mumford, D. (2003). Hierarchical bayesian inference in the visual cortex. Journal of Optical Society of America, A, 20(7), 1434–1448.
- * [1] Schölkopf, B., Burges, C. J. C., & Smola, A. J. (1999a). Advances in Kernel Methods — Support Vector Learning. MIT Press, Cambridge, MA.
- * [2] Roweis, S., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. Science, 290(5500), 2323–2326.
- *
- * [3] Bengio, Y., Ducharme, R., & Vincent, P. (2001). A neural probabilistic language model. In Leen, T., Dietterich, T., & Tresp, V. (Eds.),
- * [4] Jordan, M. I. (1998). Learning in Graphical Models. Kluwer, Dordrecht, Netherlands.
- *
- * [5] Larochelle, H., & Bengio, Y. (2008). Classification using discriminative restricted Boltzmann machines. In Cohen, W. W., McCallum, A., & Roweis, S. T. (Eds.),
- * [6] Serre, T., Kreiman, G., Kouh, M., Cadieu, C., Knoblich, U., & Poggio, T. (2007). A quantitative theory of immediate visual recognition. Progress in Brain Research, Computational neuroscience: Theoretical Insights into Brain Function, 165, 33–56.

ممنون از توجه تان