یادگیری عمیق

جلسه ۴

مبانی یادگیری عمیق

Foundations of Deep Learning

کاظم فولادی قلعه

دانشکده مهندسی، دانشکدگان فارابی

دانشگاه تهران

http://courses.fouladi.ir/deep

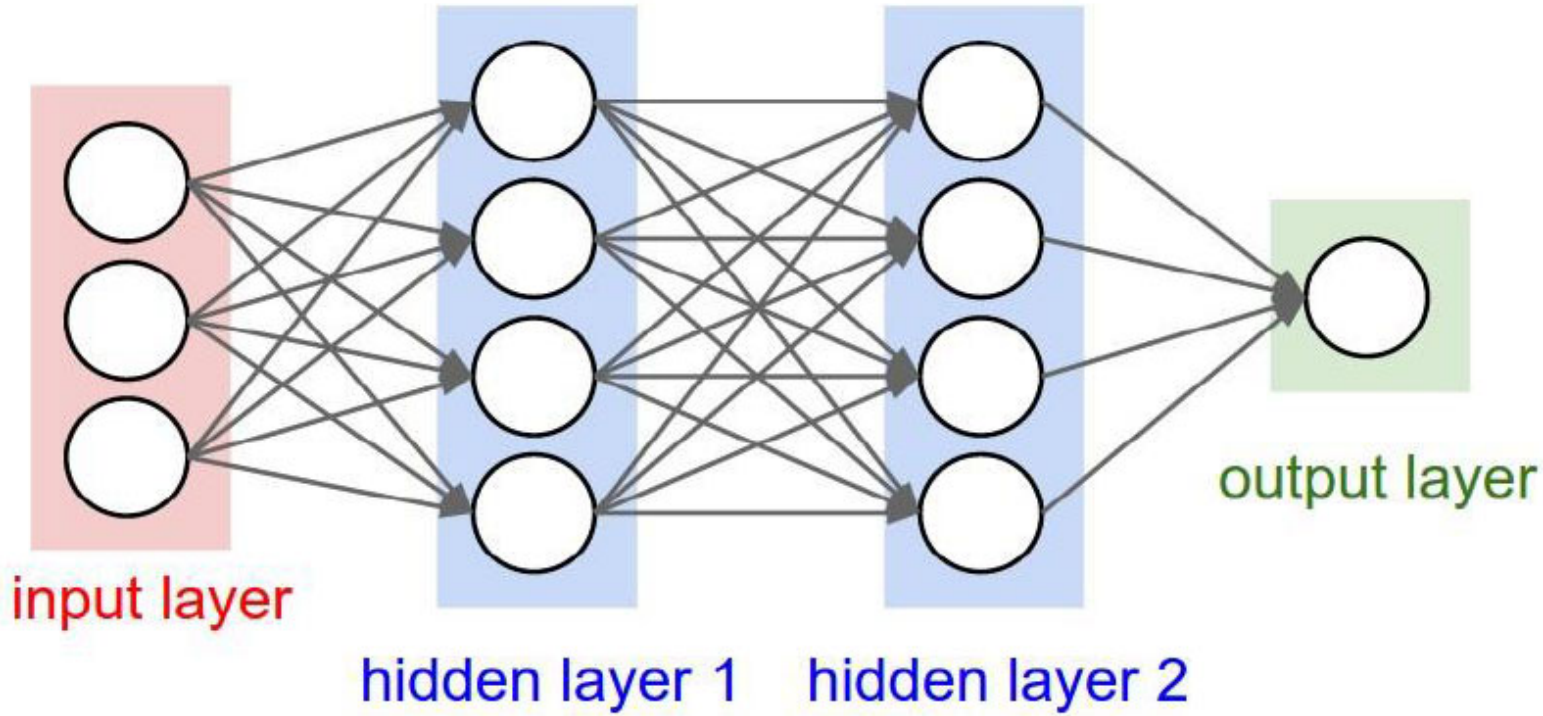
مبانی یادگیری عمیق

۱

مقدمات

Deep Learning - Basics

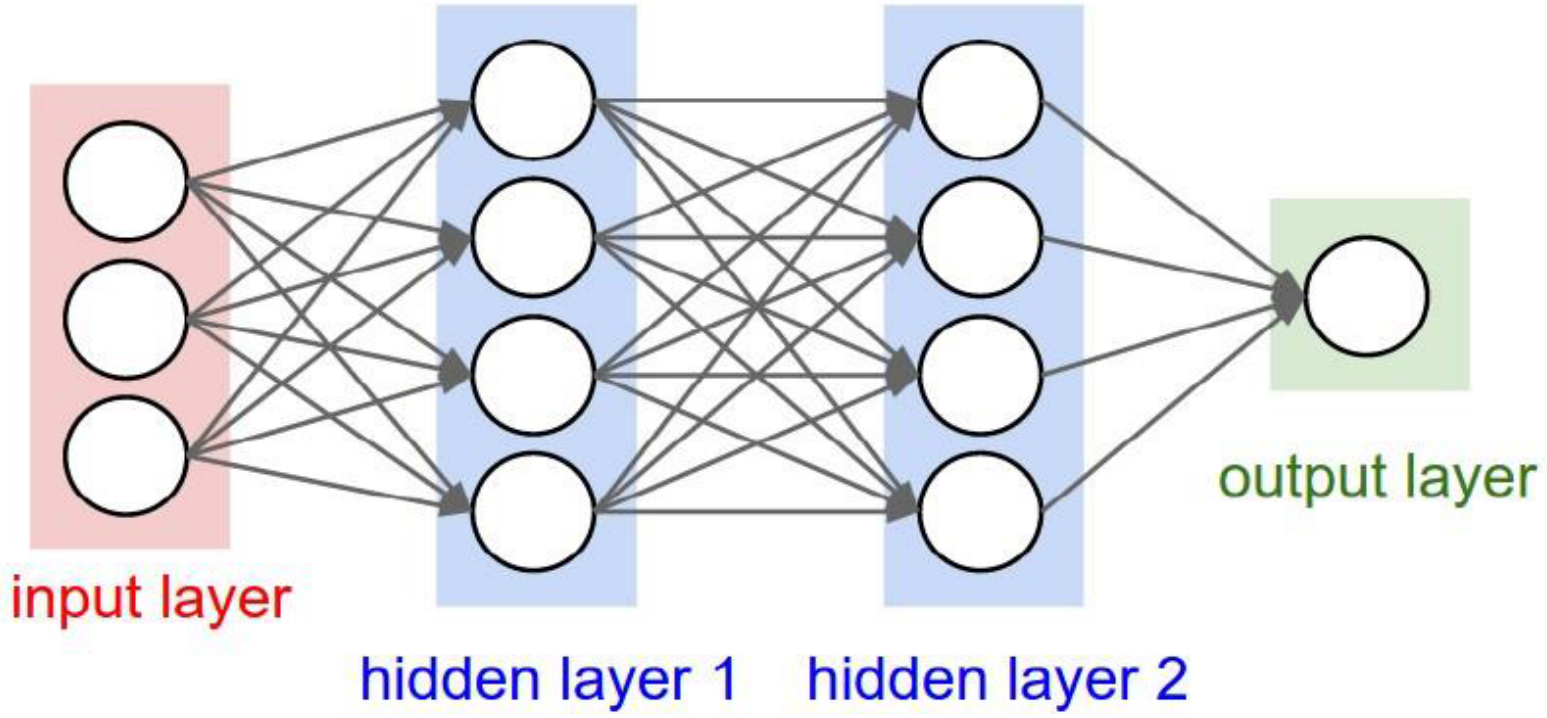
Artificial Neural Networks



Consists of one input, one output and multiple fully-connected hidden layers in-between. Each layer is represented as a series of neurons and progressively extracts higher and higher-level features of the input until the final layer essentially makes a decision about what the input shows.

شبکه‌های عصبی مصنوعی

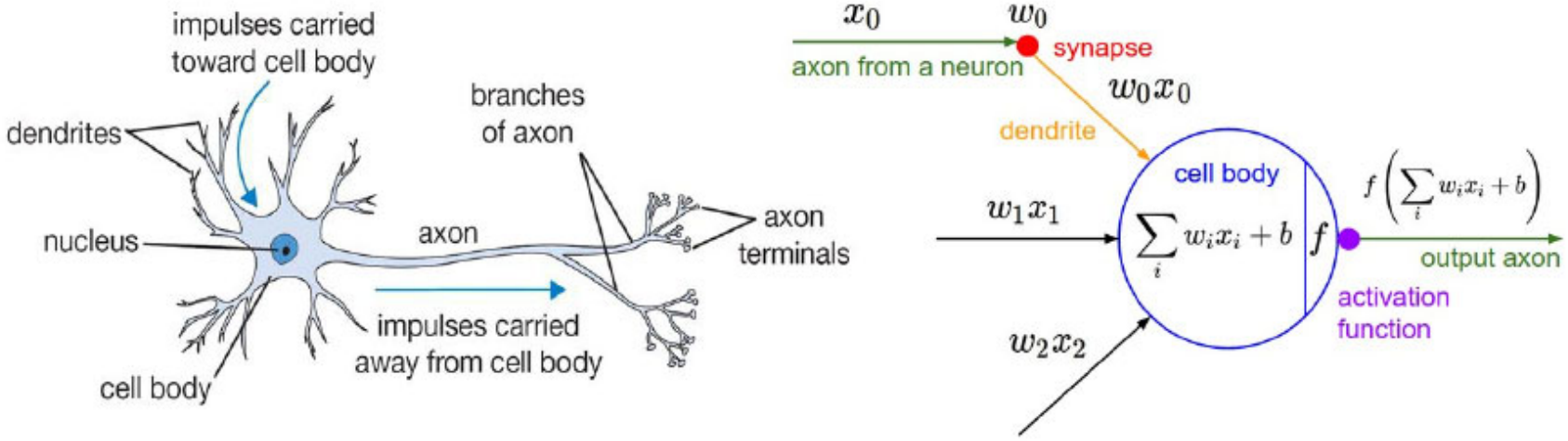
ARTIFICIAL NEURAL NETWORKS



یک شبکه‌ی عصبی (چندلایه‌ی پیش‌رو)
از یک لایه‌ی ورودی، یک لایه‌ی خروجی و چندین لایه‌ی پنهان تماماً متصل در میان آن دو تشکیل می‌شود.
هر لایه به‌صورت یک سری از نرون‌ها بازنمایی می‌شود
و به‌صورت پیش‌رونده ویژگی‌های سطح بالا و سطح بالاتری از ورودی را استخراج می‌کند
تا اینکه لایه‌ی نهایی اساساً یک تصمیم در مورد اینکه ورودی چه چیزی را نشان می‌دهد، بگیرد.

Deep Learning - Basics

The Neuron



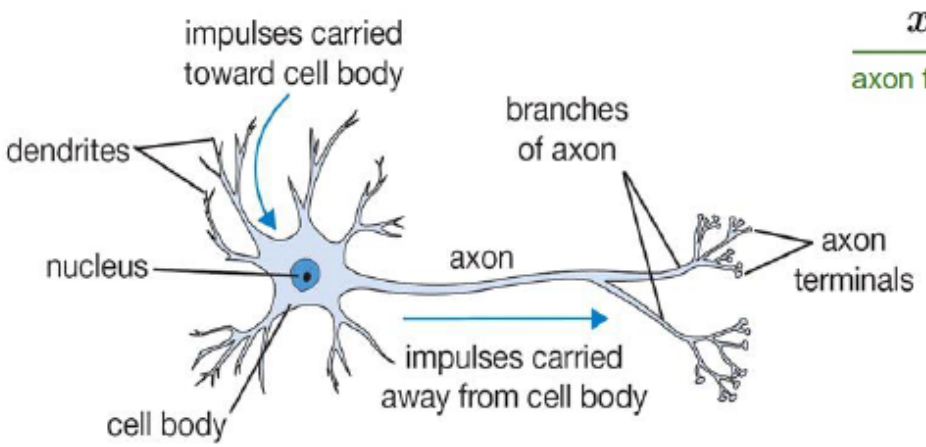
An artificial neuron contains a nonlinear activation function and has several incoming and outgoing weighted connections.

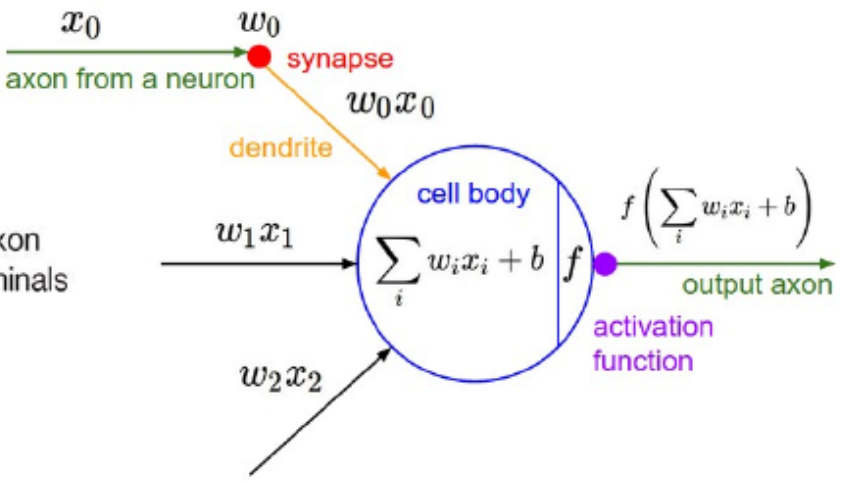
Neurons are trained to filter and detect specific features or patterns (e.g. edge, nose) by receiving weighted input, transforming it with the activation function und passing it to the outgoing connections.

شبکه‌های عصبی مصنوعی

نرون

THE NEURON



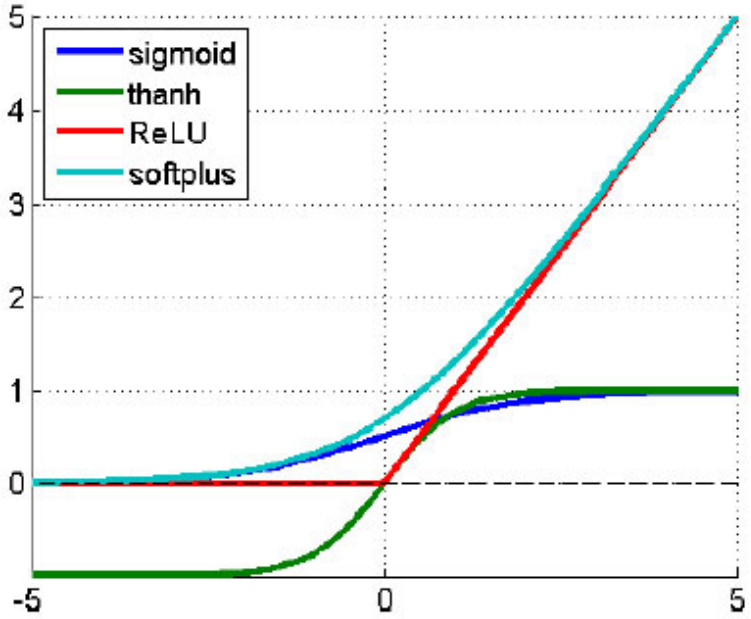


یک نرون مصنوعی، حاوی یک تابع فعال‌سازی غیرخطی است
و دارای اتصالات وزن‌دار متعدد وارد شونده و خارج شونده است.

نرون‌ها برای فیلتر کردن یا آشکار کردن ویژگی‌های مختلف آموزش می‌بینند.
این کار با تنظیم وزن‌های نرون انجام می‌شود.

Deep Learning - Basics

Non-linear Activation Function



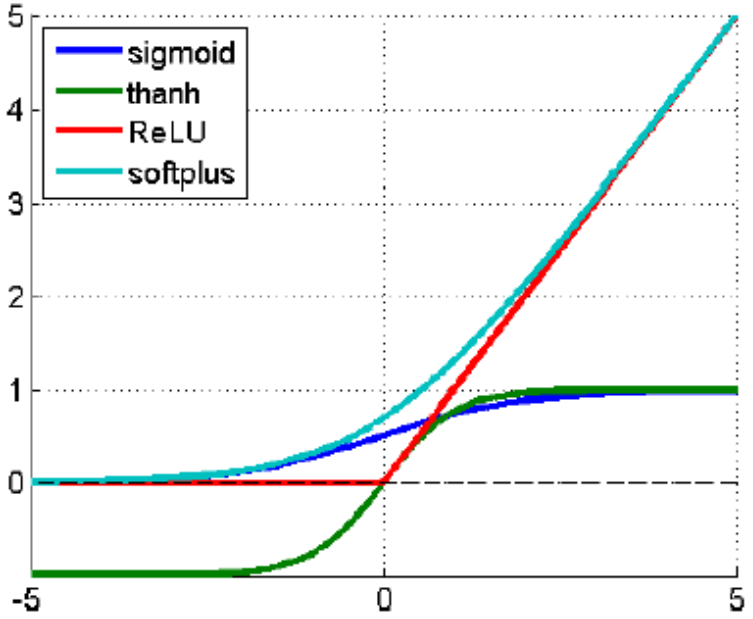
Most deep networks use ReLU - max(0,x) - nowadays for hidden layers, since it trains much faster, is more expressive than logistic function and prevents the gradient vanishing problem.

Non-linearity is needed to learn complex (non-linear) representations of data, otherwise the NN would be just a linear function.

شبکه‌های عصبی مصنوعی

تابع فعال‌سازی غیرخطی

NON-LINEAR ACTIVATION FUNCTION



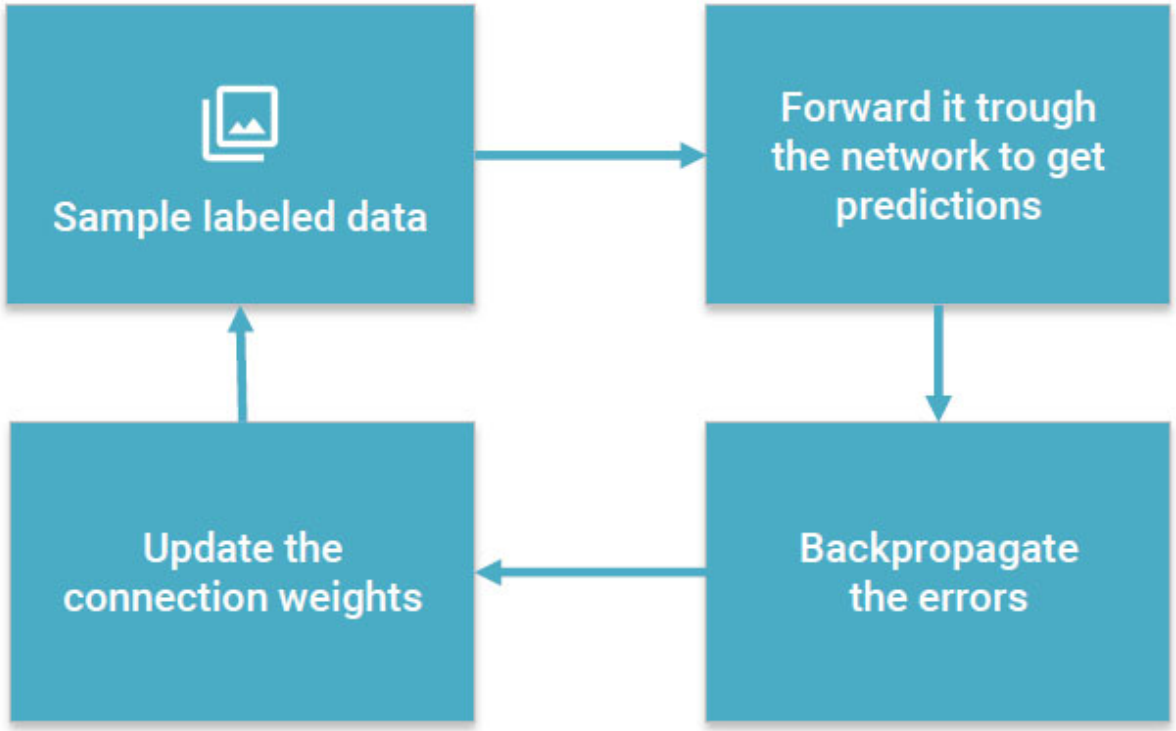
امروزه بیشتر شبکه‌های عمیق از تابع فعال‌سازی ReLU (یعنی $\max(0,x)$) برای لایه‌های پنهان استفاده می‌کنند.

چرا که فرآیند آموزش با این تابع بسیار سریع‌تر است، نسبت به تابع سیگموئید لگاریتمی رساتر است و از مشکل اضمحلال گرادیان اجتناب می‌کند.

غیرخطی بودن برای یادگیری بازنمایی‌های پیچیده (غیرخطی) داده‌ها لازم است، در غیر این صورت شبکه‌ی عصبی فقط یک تابع خطی خواهد بود.

Deep Learning - Basics

The Training Process

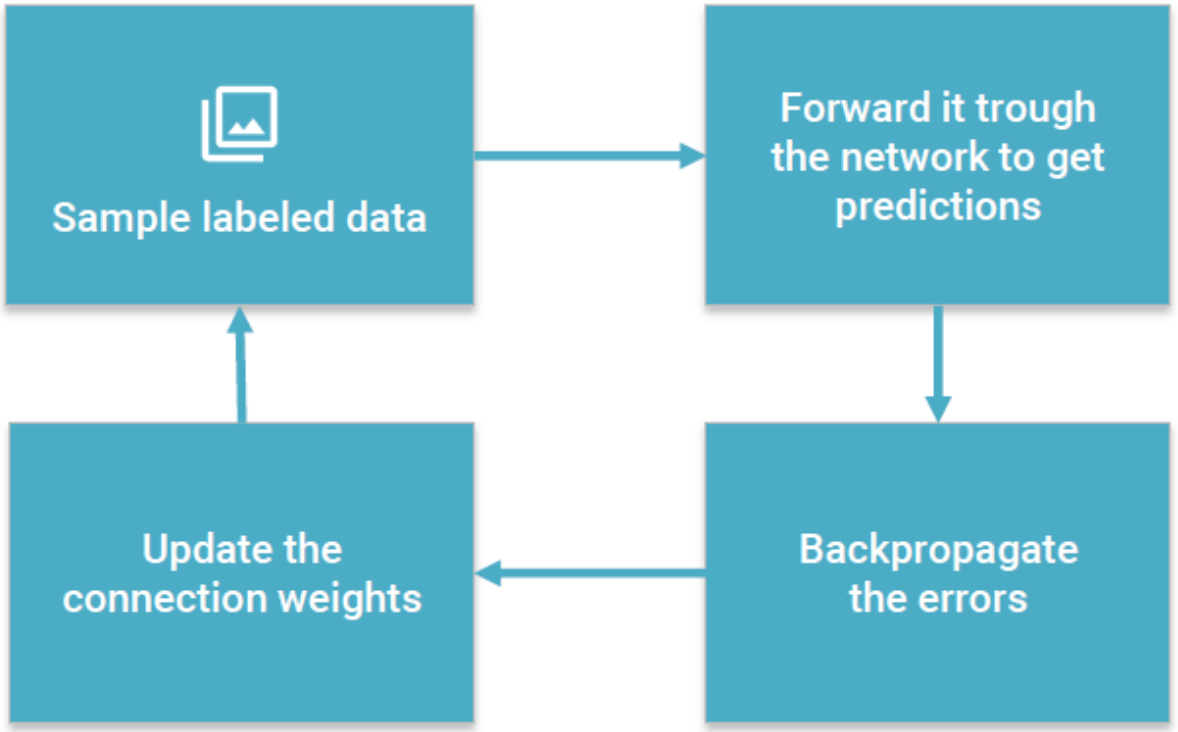


Learns by generating an error signal that measures the difference between the predictions of the network and the desired values and then using this error signal to change the weights (or parameters) so that predictions get more accurate.

شبکه‌های عصبی مصنوعی

فرآیند آموزش

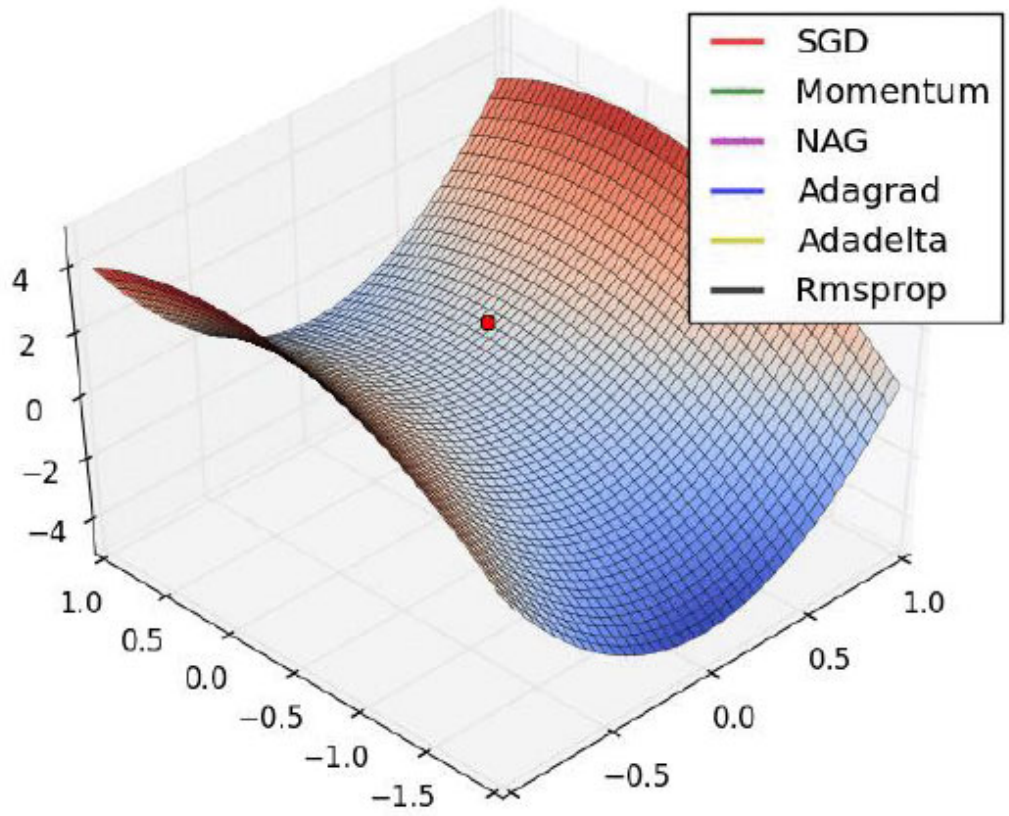
THE TRAINING PROCESS



شبکه‌ی عصبی با تولید یک سیگنال خطا که تفاوت میان پیش‌بینی‌های شبکه و مقادیر مطلوب را اندازه‌گیری می‌کند و سپس استفاده از این سیگنال خطا برای تغییر وزن‌ها (یا پارامترها) به‌منظور دقیق‌تر کردن پیش‌بینی‌ها یاد می‌گیرد.

Deep Learning - Basics

Gradient Descent

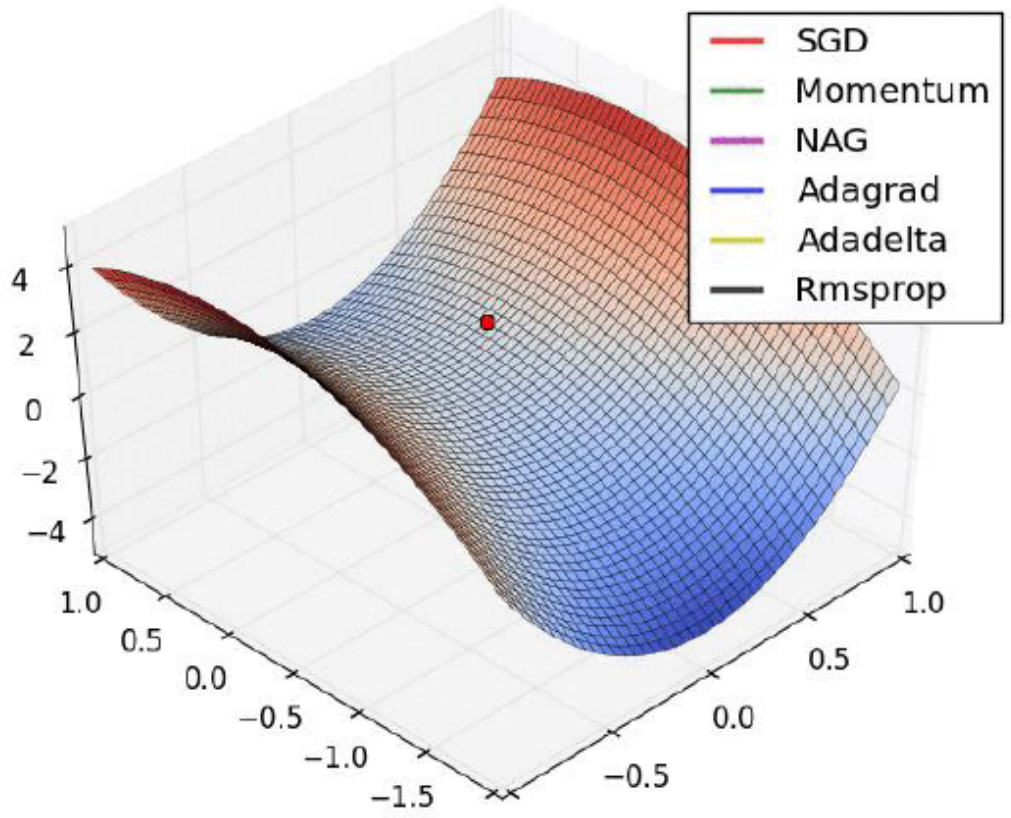


Gradient Descent finds the (local) the minimum of the cost function (used to calculate the output error) and is used to adjust the weights.

شبکه‌های عصبی مصنوعی

کاهش گرادیان

GRADIENT DESCENT



الگوریتم بهینه‌سازی کاهش گرادیان می‌نیمم (محلی) تابع هزینه (که خطای خروجی را حساب می‌کند) را می‌یابد. این الگوریتم برای تنظیم وزن‌ها استفاده می‌شود.

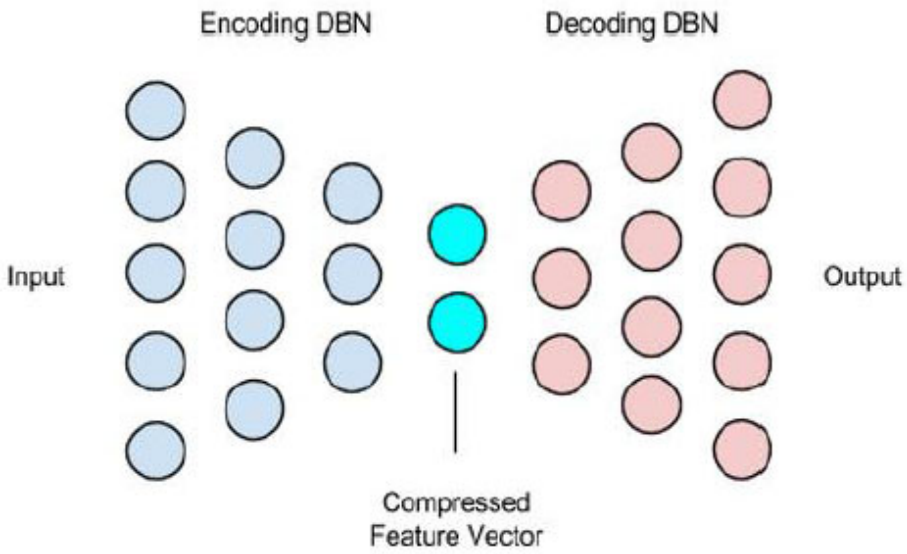
مبانی یادگیری عمیق

۲

نمونه‌هایی از معماری‌های عمیق

Deep Learning - Basics

Deep Autoencoders

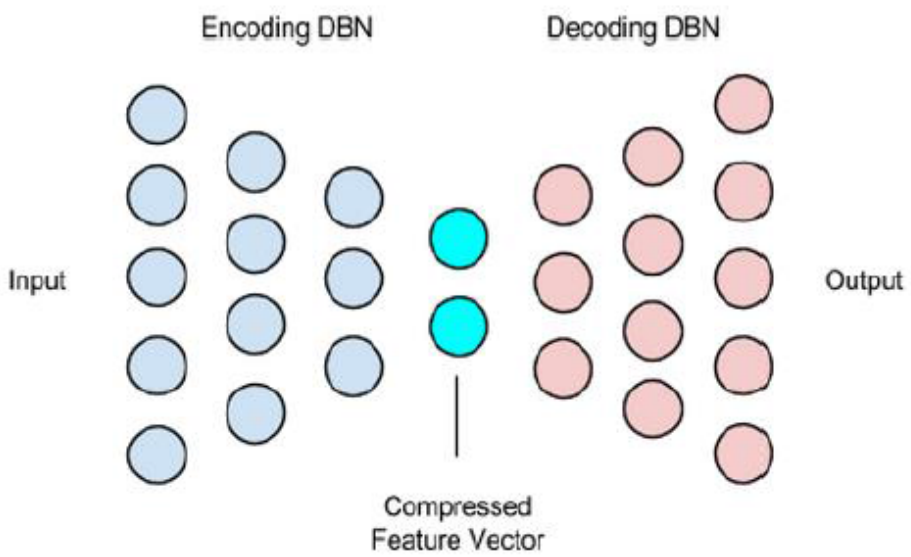


Composed of two symmetrical deep-belief networks. The encoding network learns to compresses the input to a condensed vector (dimensionality reduction). The decoding network can be used to reconstruct the data.

Topic Modeling: Document in a collection is converted to a Bag-of-Words and transformed to a compressed feature vector using an autoencoder. The distance from every other document-vector can be measured and nearby document-vectors fall under the same topic.

خودکدگذارهای عمیق

DEEP AUTOENCODERS



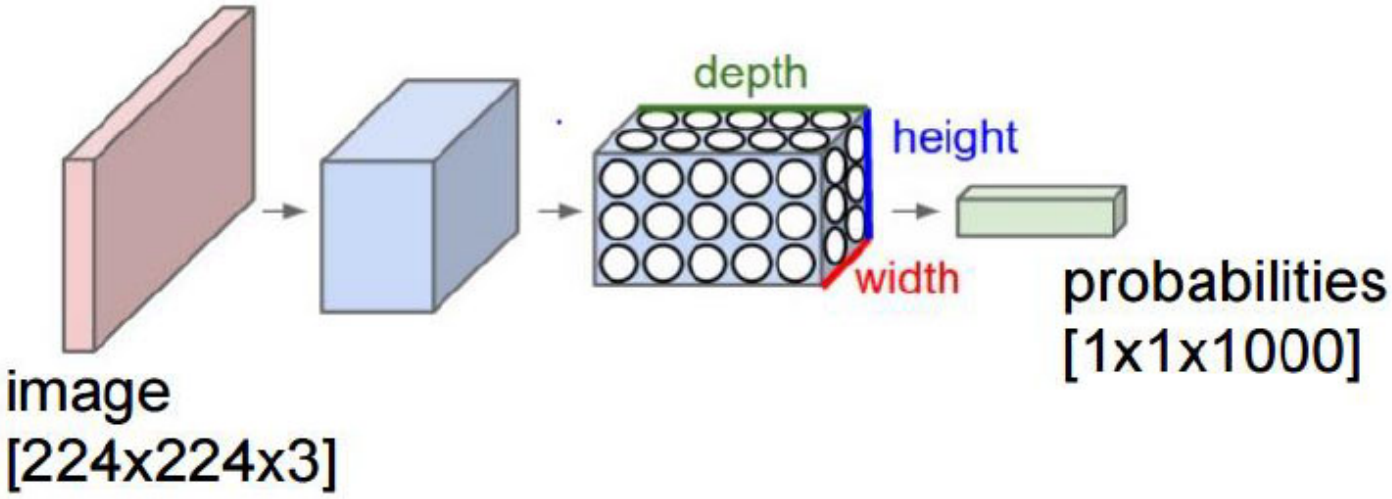
خودکدگذار عمیق از دو شبکه‌ی باور عمیق متقارن تشکیل می‌شود.
شبکه‌ی کدگذاری یاد می‌گیرد که ورودی را در یک بردار متراکم فشرده کند. (کاهش بعد)
شبکه‌ی کدگشایی می‌تواند برای بازسازی داده‌ها استفاده شود.

مثال از کاربرد: **مدل‌سازی موضوع (Topic Modeling)**
یک سند در یک گردایه، به کیسه‌ای-از-کلمات (Bag-of-Words) تبدیل می‌شود و سپس با استفاده از یک خودکدگذار به یک بردار ویژگی فشرده شده تبدیل می‌شود.
فاصله از هر سند-بردار دیگر می‌تواند اندازه‌گیری شود
و سند-بردارهای نزدیک به‌هم تحت یک موضوع یکسان قرار می‌گیرند.

Deep Learning - Basics

Convolutional Neural Nets (CNN)

Convolutional Neural Networks learn a complex representation of visual data using vast amounts of data. They are inspired by the human visual system and learn multiple layers of transformations, which are applied on top of each other to extract a progressively more sophisticated representation of the input.

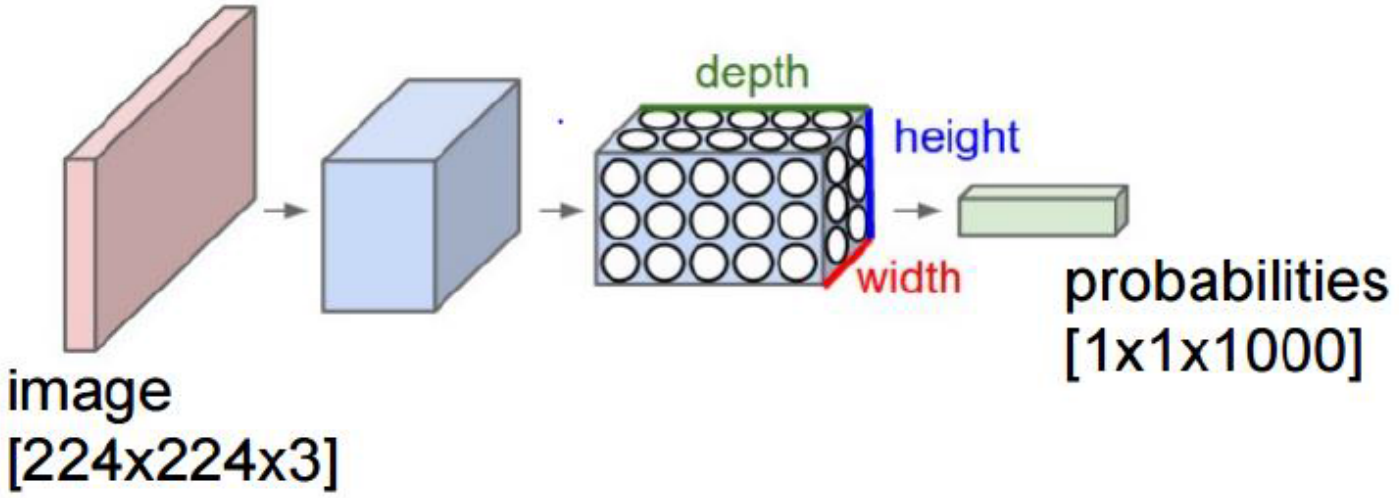


Every layer of a CNN takes a 3D volume of numbers and outputs a 3D volume of numbers. E.g. Image is a 224*224*3 (RGB) cube and will be transformed to 1*1000 vector of probabilities.

شبکه‌های عصبی کانوولوشنال

CONVOLUTIONAL NEURAL NETWORKS (CNN)

شبکه‌های عصبی کانوولوشنال، یک بازنمایی پیچیده از داده‌های دیداری را
با استفاده از مقدار حجیمی از داده‌ها یاد می‌گیرند.
این شبکه‌ها از سیستم بینایی انسان الهام گرفته شده‌اند و چندین لایه از تبدیلات را یاد می‌گیرند،
که هر یک بر روی دیگری اعمال می‌شود و به‌صورت پیش‌رونده،
بازنمایی پیچیده‌تری از ورودی را استخراج می‌کنند.

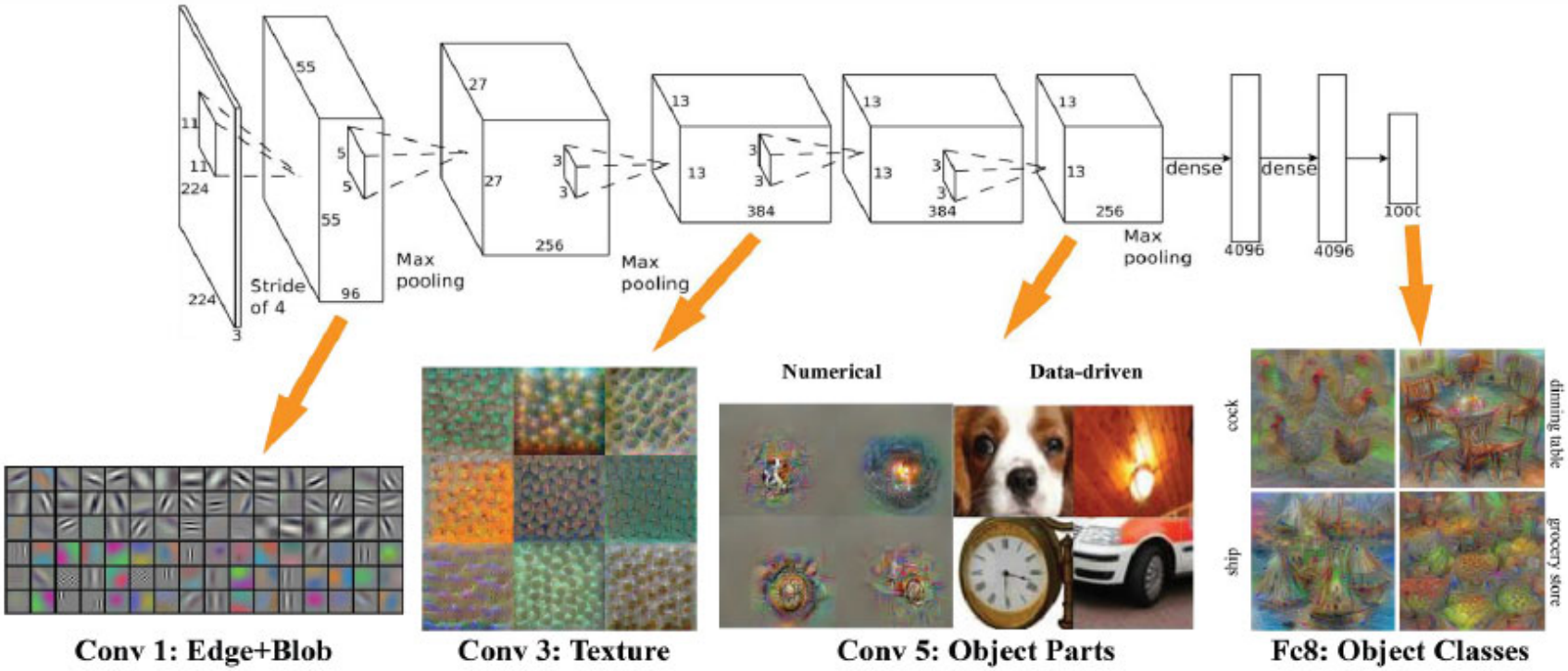


مثال از کاربرد: **دسته‌بندی تصاویر (Image Categorization)**

هر لایه از یک CNN یک حجم سه–بعدی از اعداد را دریافت می‌کند و یک حجم سه–بعدی از اعداد را خروجی می‌دهد.
برای مثال: تصویر یک مکعب $3 \times 224 \times 224$ (RGB) است و به یک بردار 1000 تایی از احتمالات تبدیل می‌شود.

Deep Learning - Basics

Convolutional Neural Nets (CNN)



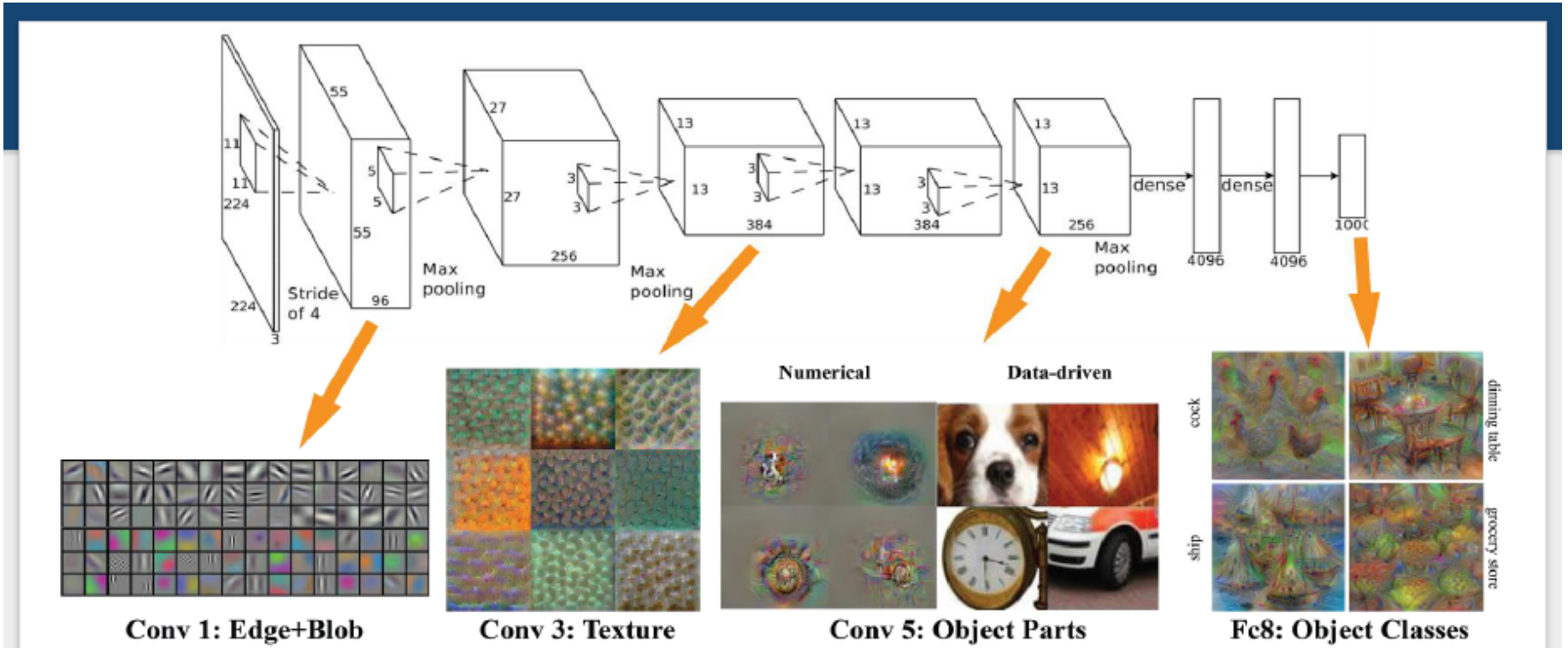
Convolution layer is a feature detector that automagically learns to filter out not needed information from an input by using convolution kernel.

Pooling layers compute the max or average value of a particular feature over a region of the input data *(downsizing of input images)*. Also helps to detect objects in some unusual places and reduces memory size.

شبکه‌های عصبی کانوولوشنال

لایه‌های اصلی

CONVOLUTIONAL NEURAL NETWORKS (CNN)



لایه‌ی کانوولوشن
Convolution Layer

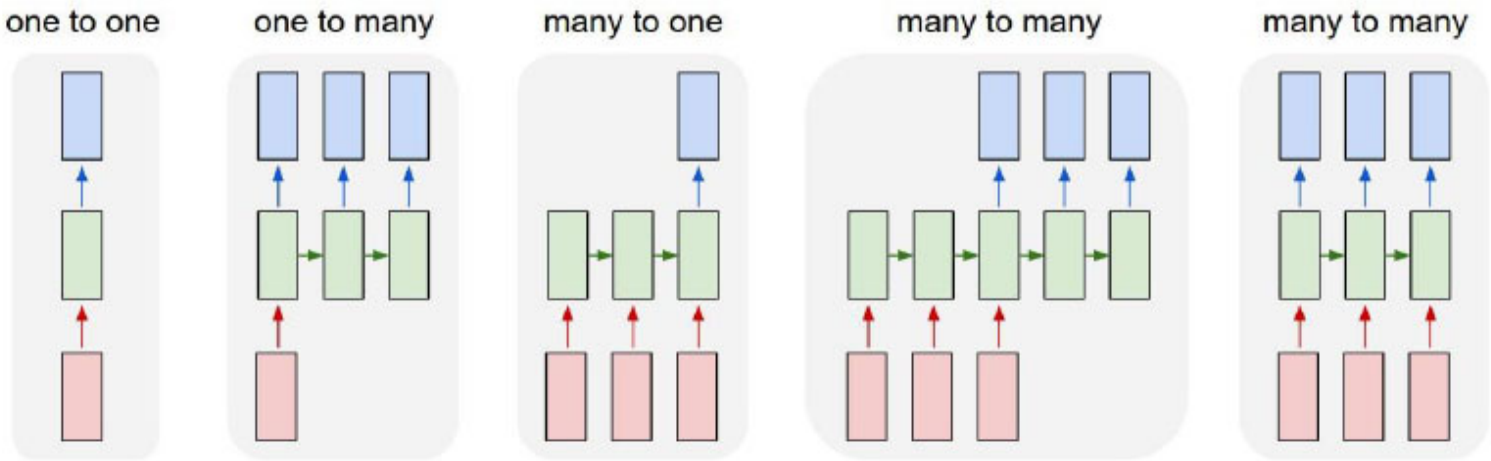
لایه‌ی کانوولوشن، یک آشکارساز ویژگی است که به‌صورت خودکار و جادویی یاد می‌گیرد که با استفاده از کرنل کانوولوشن اطلاعات غیر لازم از یک ورودی را فیلتر کند .

لایه‌ی تلفیق
Pooling Layer

لایه‌ی تلفیق، ماکزیمم یا مقدار متوسط ویژگی‌های خاص بر روی یک ناحیه از داده‌های ورودی را محاسبه می‌کند (⇐ کاهش اندازه‌ی تصاویر ورودی).
همچنین به آشکارسازی اشیا در برخی مکان‌های نامعمول کمک می‌کند و اندازه‌ی حافظه را کاهش می‌دهد.

Deep Learning - Basics

Recurrent Neural Nets (RNN)

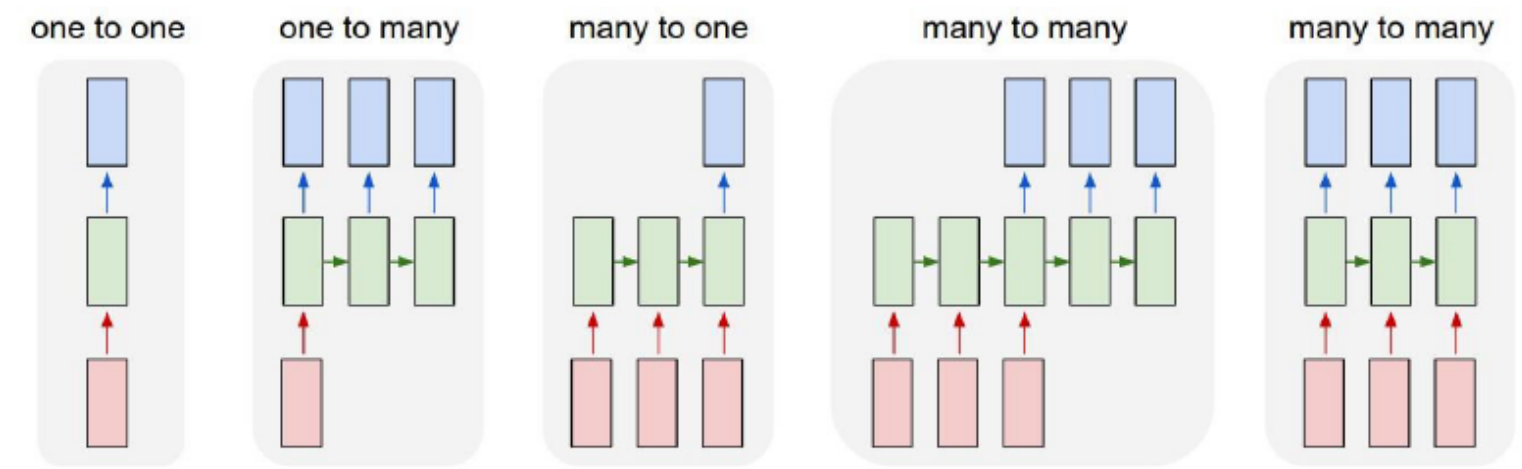


RNNs are general computers which can learn algorithms to map input sequences to output sequences (flexible-sized vectors). The output vector's contents are influenced by the entire history of inputs.

State-of-the-art results in time series prediction, adaptive robotics, handwriting recognition, image classification, speech recognition, stock market prediction, and other sequence learning problems. *Everything can be processed sequentially.*

شبکه‌های عصبی بازگشتی

Recurrent Neural Networks (RNN)



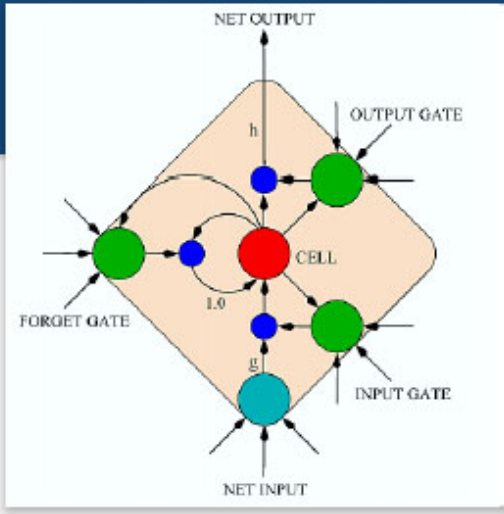
شبکه‌های عصبی بازگشتی، مدل‌های محاسباتی عمومی هستند که می‌توانند الگوریتم‌هایی را برای نگاشت **دنباله‌های ورودی** به **دنباله‌های خروجی** را یاد بگیرند.
محتوای بردار خروجی، از کل تاریخچه‌ی ورودی‌ها متأثر می‌شود.

نتایج مرز دانشی در پیش‌بینی سری زمانی، رباتیک وفقی، بازشناسی دست‌نویس، طبقه‌بندی تصاویر، بازشناسی گفتار، پیش‌بینی بازار سهام و دیگر مسائل یادگیری دنباله، متعلق به شبکه‌های عصبی بازگشتی (RNN) است.
* ... و هر چیزی که بتواند به‌صورت دنباله‌ای/ترتیبی پردازش شود.

http://colah.github.io/posts/2015-08-Understanding-LSTMs/

Deep Learning - Basics

Long Short-Term Memory RNN (LSTM)



A Long Short-Term Memory (LSTM) network is a particular type of recurrent network that works slightly better in practice, owing to its more powerful update equation and some appealing back propagation dynamics.

The **LSTM units** give the network memory cells with read, write and reset operations. During training, the network can learn when it should remember data and when it should throw it away.

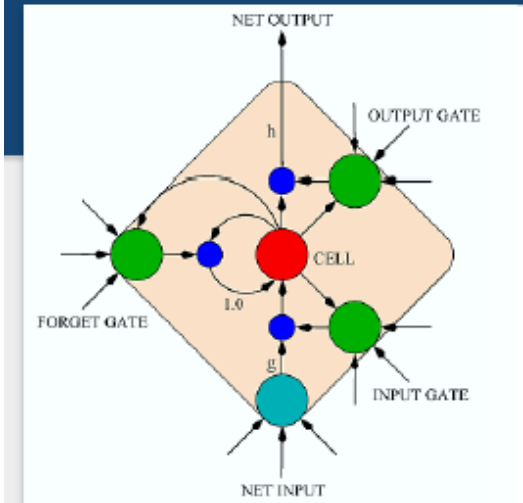
Well-suited to learn from experience to classify, process and predict time series when there are very long time lags of unknown size between important events.

Source

شبکه‌های عصبی بازگشتی

شبکه‌های عصبی بازگشتی حافظه‌ی کوتاه‌مدت طولانی

LONG SHORT-TERM MEMORY RNN (LSTM)



یک شبکه‌ی حافظه‌ی کوتاه‌مدت طولانی (LSTM)
یک نوع خاص از شبکه‌های بازگشتی است که
در عمل کمی بهتر کار می‌کند،
به‌دلیل: معادله‌ی به‌هنگام‌سازی قدرت‌مندتر آن و
چند دینامیک جالب پس‌انتشار در آن

واحدهای LSTM، به سلول‌های حافظه‌ی شبکه،
امکان عملیات خواندن (read)، نوشتن (write) و بازنشانی (reset) را می‌دهند.
در حین آموزش، شبکه می‌تواند یاد بگیرد که
چه زمانی باید داده‌ها را به‌خاطر بیاورد و چه زمانی باید آن را دور بریزد.

شبکه‌ی LSTM، برای یادگیری از تجربه در طبقه‌بندی، پردازش و پیش‌بینی سری‌های زمانی،
مناسب است هرگاه زمان‌های عقب‌ماندگی (lag) بسیار طولانی با اندازه‌ی نامعلوم بین رویدادهای
مهم وجود داشته باشد.

http://www.wildml.com/2016/01/attention-and-memory-in-deep-learning-and-nlp/

Deep Learning - Basics

Recurrent Neural Nets (RNN) – Attention Mechanisms

by *ent423* , *ent261* correspondent updated 9:49 pm et , thu march 19 , 2015 (*ent261*) a *ent114* was killed in a parachute accident in *ent45* , *ent85* , near *ent312* , a *ent119* official told *ent261* on wednesday . he was identified thursday as special warfare operator 3rd class *ent23* , 29 , of *ent187* , *ent265* . `` *ent23* distinguished himself consistently throughout his career . he was the epitome of the quiet professional in all facets of his life , and he leaves an inspiring legacy of natural tenacity and focused

. . .

ent119 identifies deceased sailor as **X** , who leaves behind a wife

by *ent270* , *ent223* updated 9:35 am et , mon march 2 , 2015 (*ent223*) *ent63* went familial for fall at its fashion show in *ent231* on sunday , dedicating its collection to `` mamma '' with nary a pair of `` mom jeans '' in sight . *ent164* and *ent21* , who are behind the *ent196* brand , sent models down the runway in decidedly feminine dresses and skirts adorned with roses , lace and even embroidered doodles by the designers ' own nieces and nephews . many of the looks featured saccharine needlework phrases like `` i love you ,

. . .

X dedicated their fall fashion show to moms

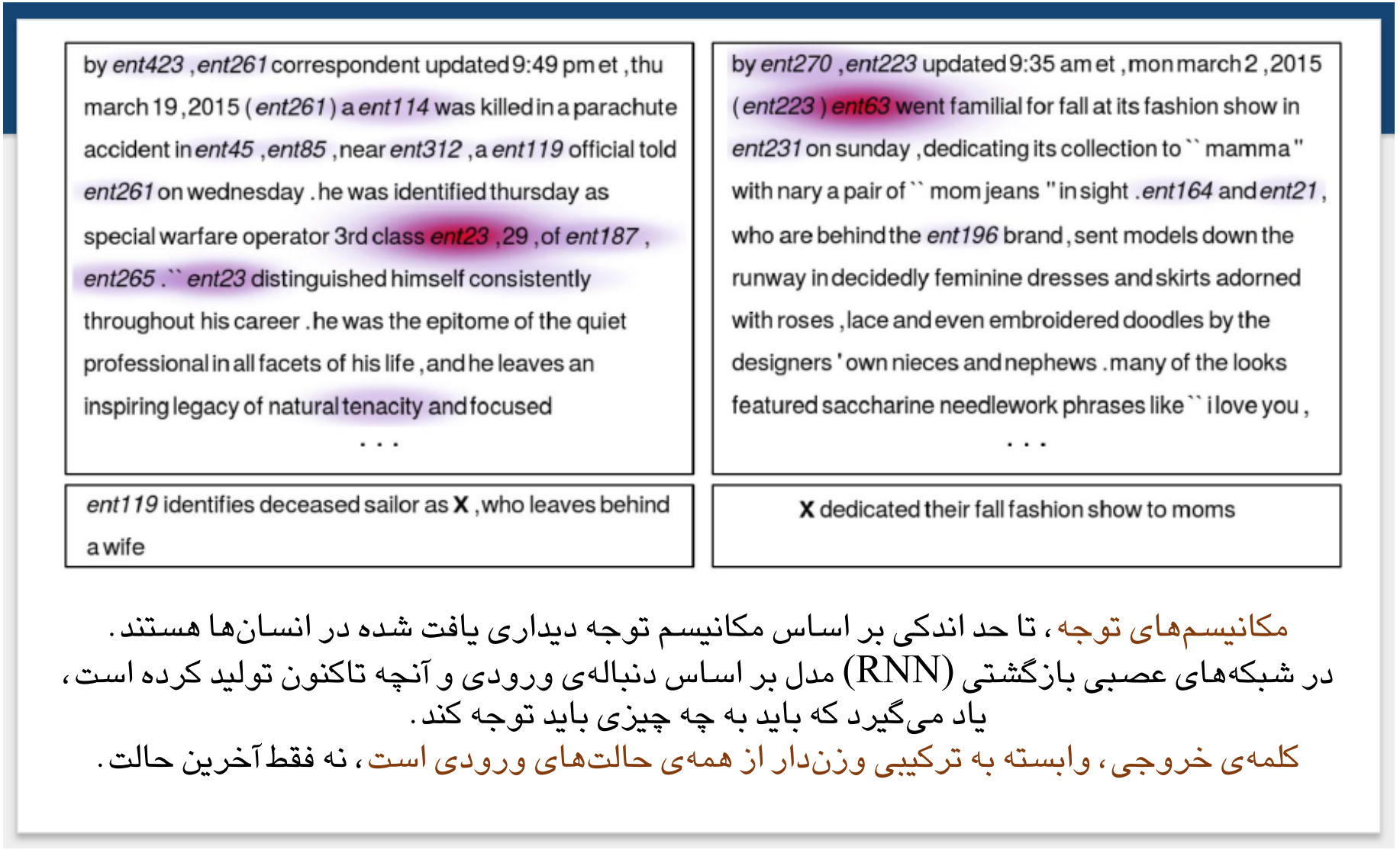
Attention Mechanisms are loosely based on the visual attention mechanism found in humans. In RNNs the model learns what to attend to based on the input sentence and what it has produced so far. The output word depends on a weighted combination of all the input states, not just the last state.

Source

شبکه‌های عصبی بازگشتی

مکانیسم‌های توجه

ATTENTION MECHANISMS

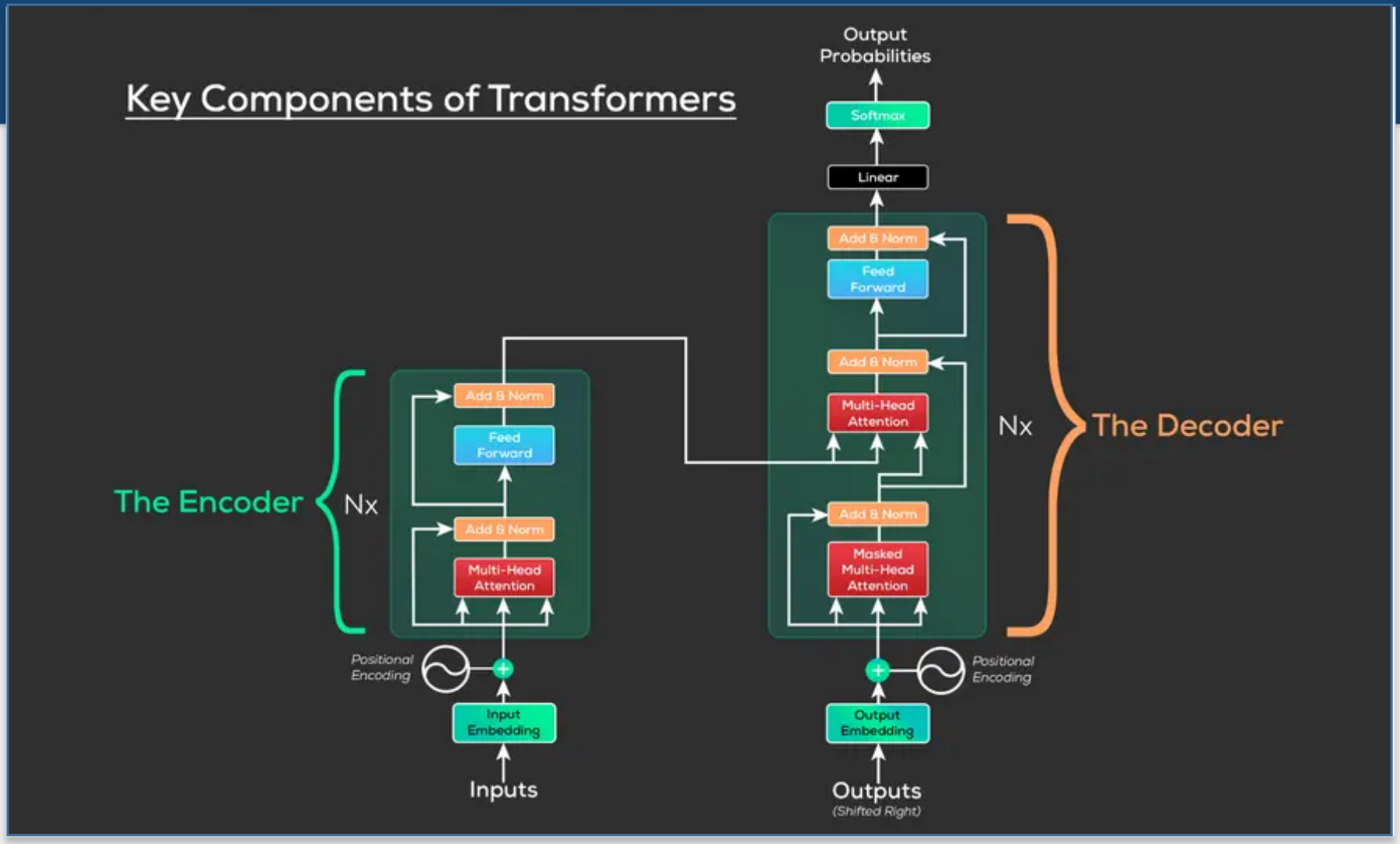


مکانیسم‌های توجه، تا حد اندکی بر اساس مکانیسم توجه دیداری یافت شده در انسان‌ها هستند.
در شبکه‌های عصبی بازگشتی (RNN) مدل بر اساس دنباله‌ی ورودی و آنچه تاکنون تولید کرده است، یاد می‌گیرد که باید به چه چیزی باید توجه کند.
کلمه‌ی خروجی، وابسته به ترکیبی وزن‌دار از همه‌ی حالت‌های ورودی است، نه فقط آخرین حالت.

https://deepgram.com/ai-glossary/transformers

Deep Learning - Basics

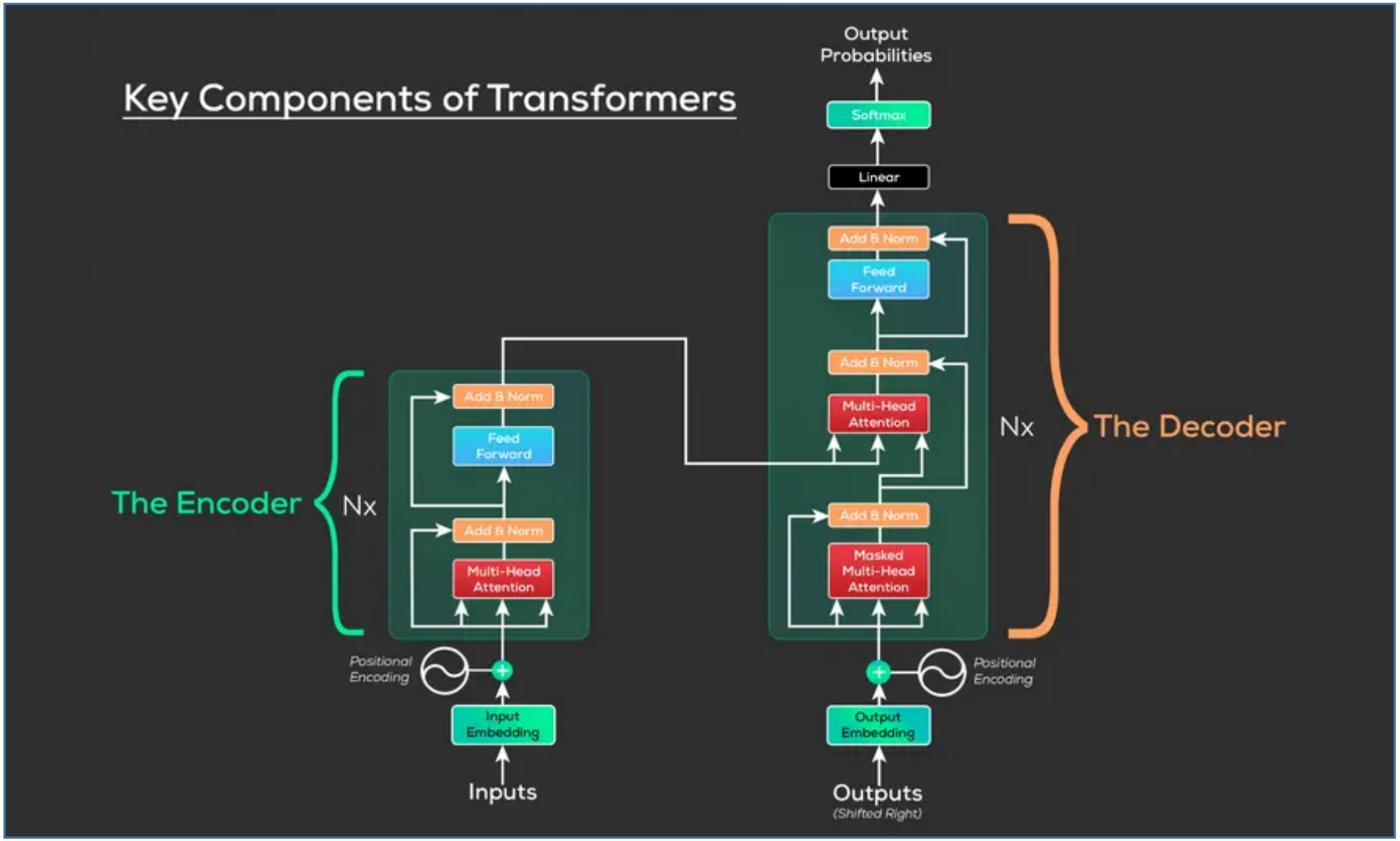
Transformers



Source

مبدل‌ها (ترانسفورمرها)

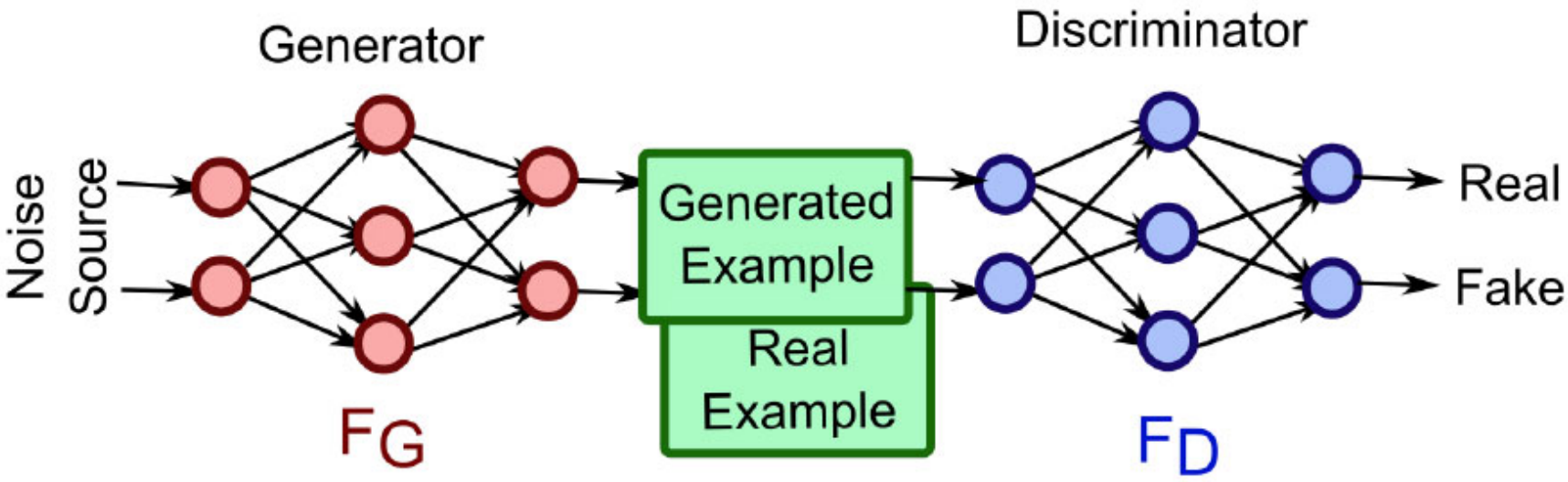
TRANSFORMERS



https://blog.openai.com/generative-models/

Deep Learning - Basics

Generative Adversarial Networks

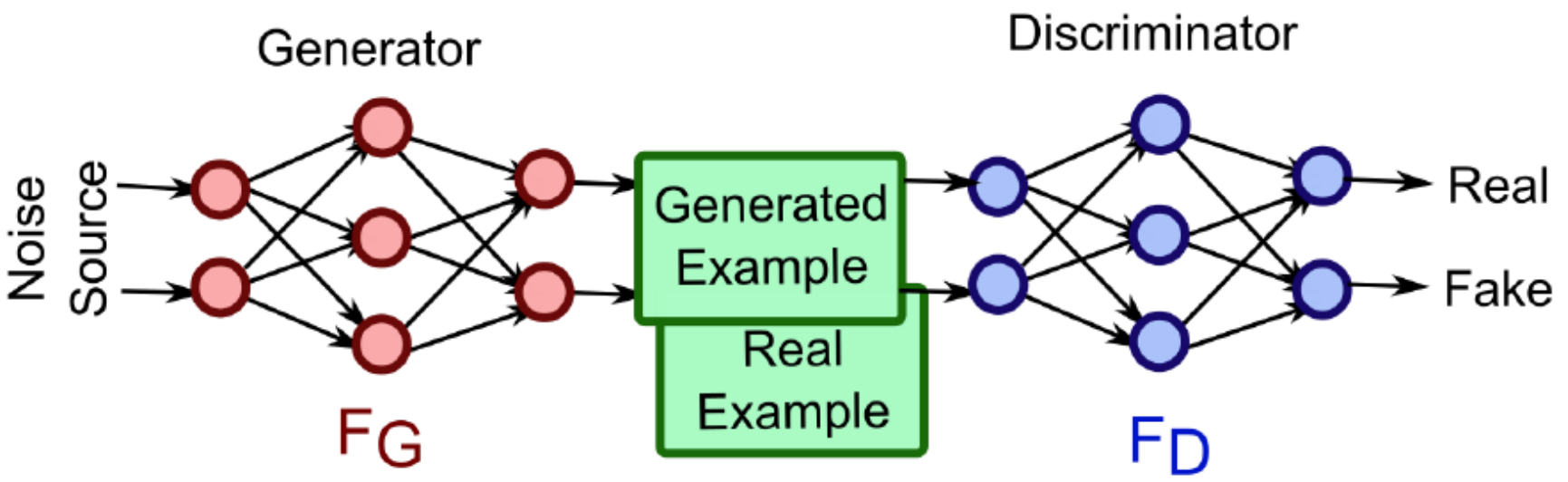


Generative Adversarial Networks (GANs) consist of any two networks with one tasked to generate content and the other has to judge content. The discriminating network receives either training data or generated content from the generative network and tries to predict the data source (real or fake). This creates a form of competition where the discriminator is getting better at distinguishing real data from generated data and the generator is learning to become less predictable to the discriminator.

Source

شبکه‌های تخاصمی مولد

GENERATIVE ADVERSARIAL NETWORKS (GAN)

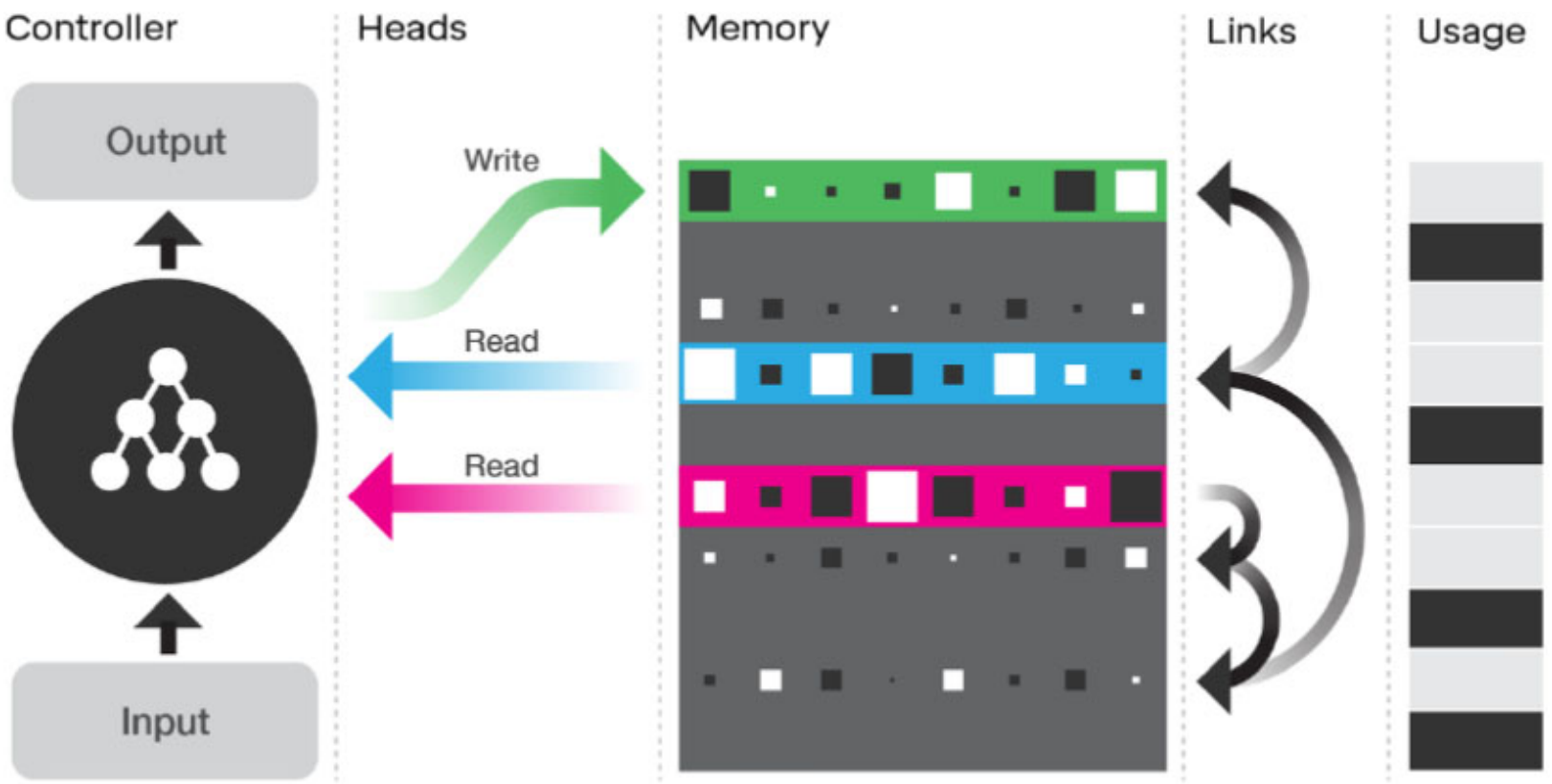


شبکه‌های تخاصمی مولد (GANs) از دو شبکه تشکیل شده است که
یکی وظیفه دارد محتوا تولید کند و دیگری باید محتوا را داوری کند.
شبکه‌ی تفکیک‌کننده، داده‌های آموزشی یا محتوای تولید شده توسط شبکه‌ی مولد را دریافت می‌کند و
سعی می‌کند منبع داده (واقعی 'real' یا تقلبی 'fake') را پیش‌بینی کند.
این شبکه نوعی رقابت ایجاد می‌کند که در آن تفکیک‌گر با تمایزبخشی داده‌های واقعی از داده‌های تولیدی،
و تولیدگر با یادگیری قابلیت پیش‌بینی کمتر برای تفکیک‌گر بهتر می‌شود.

https://deepmind.com/blog/differentiable-neural-computers/

Deep Learning - Basics

Differentiable Neural Computer

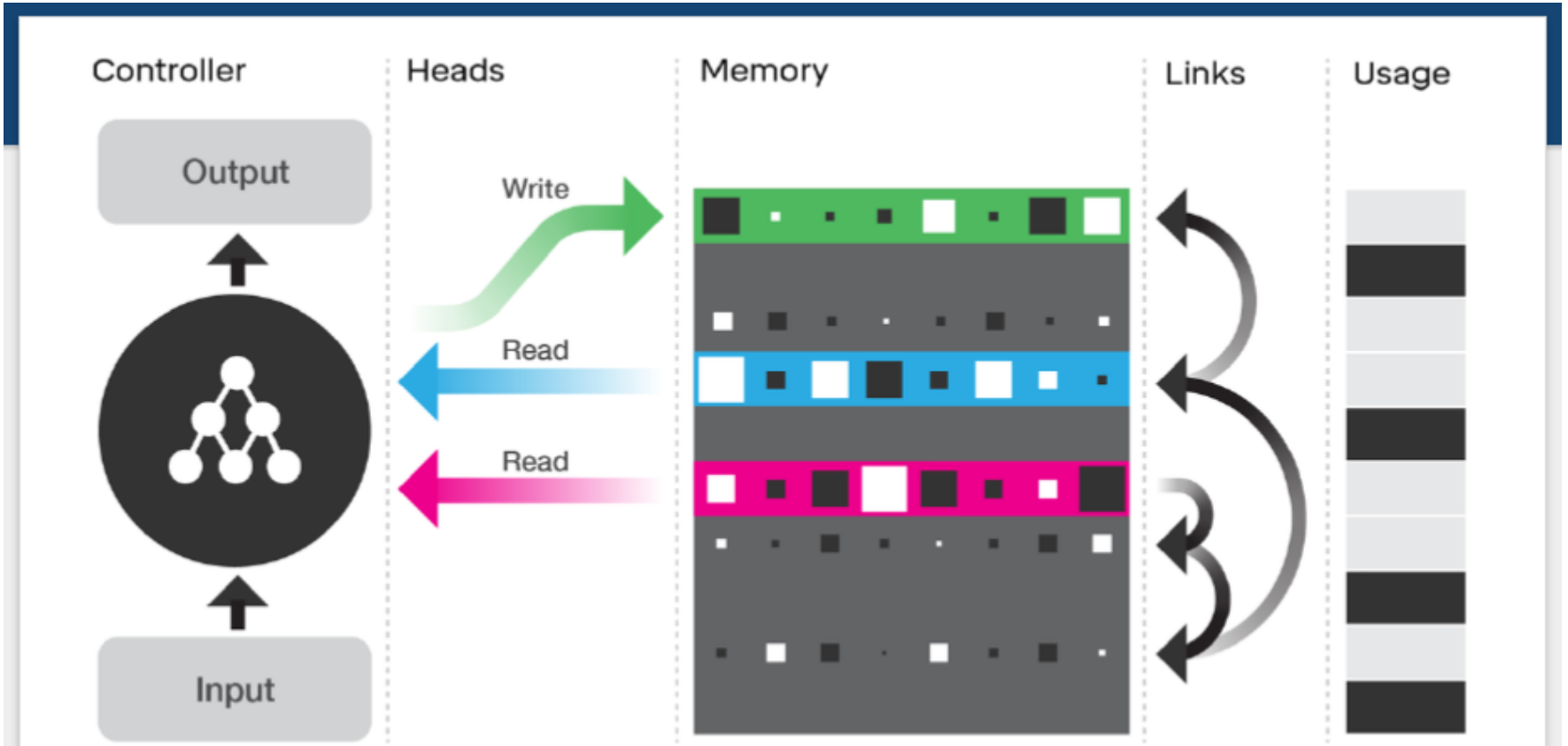


Differentiable Neural Computer is a hybrid learning machine combining neural networks with read-write memory. They learn how to use memory and how to produce answers completely from scratch. This learning machine is able, without prior programming, to organize information into connected facts and use those facts to solve problems.

Source

کامپیوتر عصبی تفاضل‌پذیر

DIFFERENTIABLE NEURAL COMPUTER

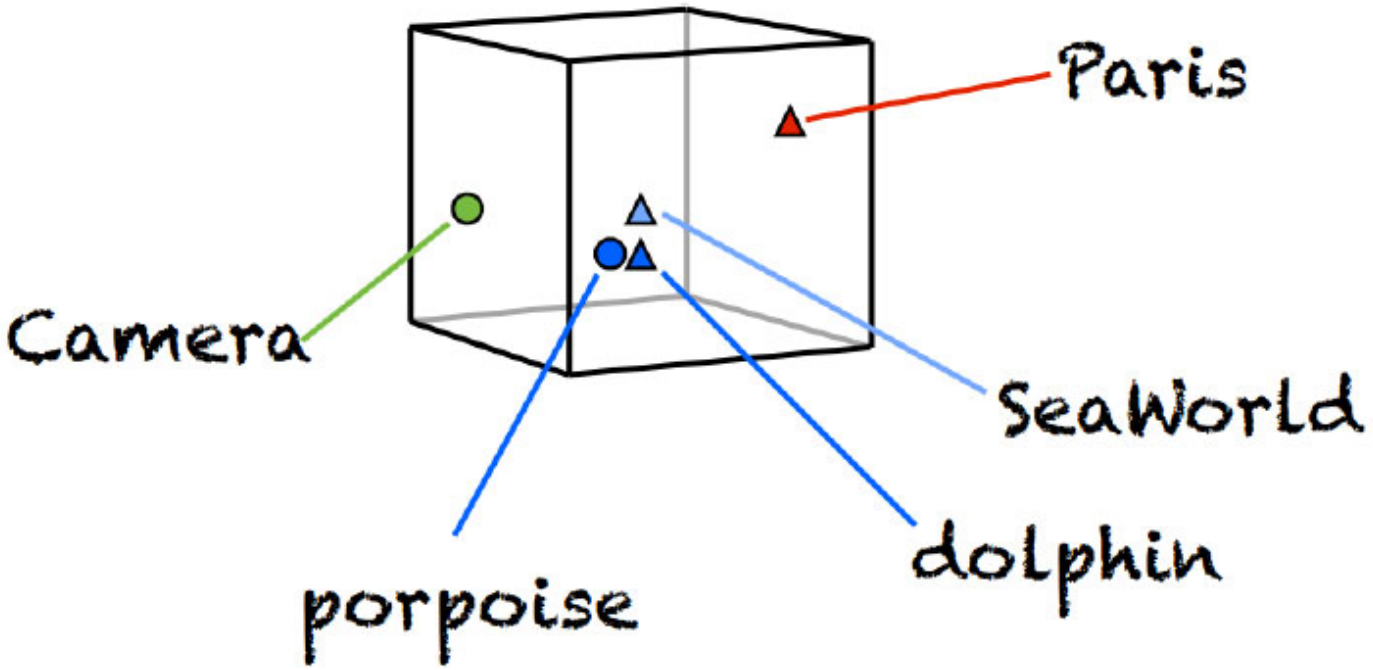


«کامپیوتر عصبی تفاضل‌پذیر» یک ماشین یادگیری هیبریدی است که
شبکه‌های عصبی را با حافظه‌ی خواندن–نوشتن ترکیب می‌کند.
این کامپیوترها یاد می‌گیرند که چگونه باید از حافظه استفاده کنند و چگونه باید پاسخ‌ها را کاملاً از صفر تولید کنند.
این ماشین یادگیری، قادر است بدون برنامه‌ریزی پیشین
اطلاعات را در فکت‌های متصل سازمان‌دهی کند و از آن فکت‌ها برای حل مسائل استفاده کند.

http://static.googleusercontent.com/media/research.google.com/en/people/jeff/CIKM-keynote-Nov2014.pdf

Deep Learning - Basics

Natural Language Processing – Embeddings



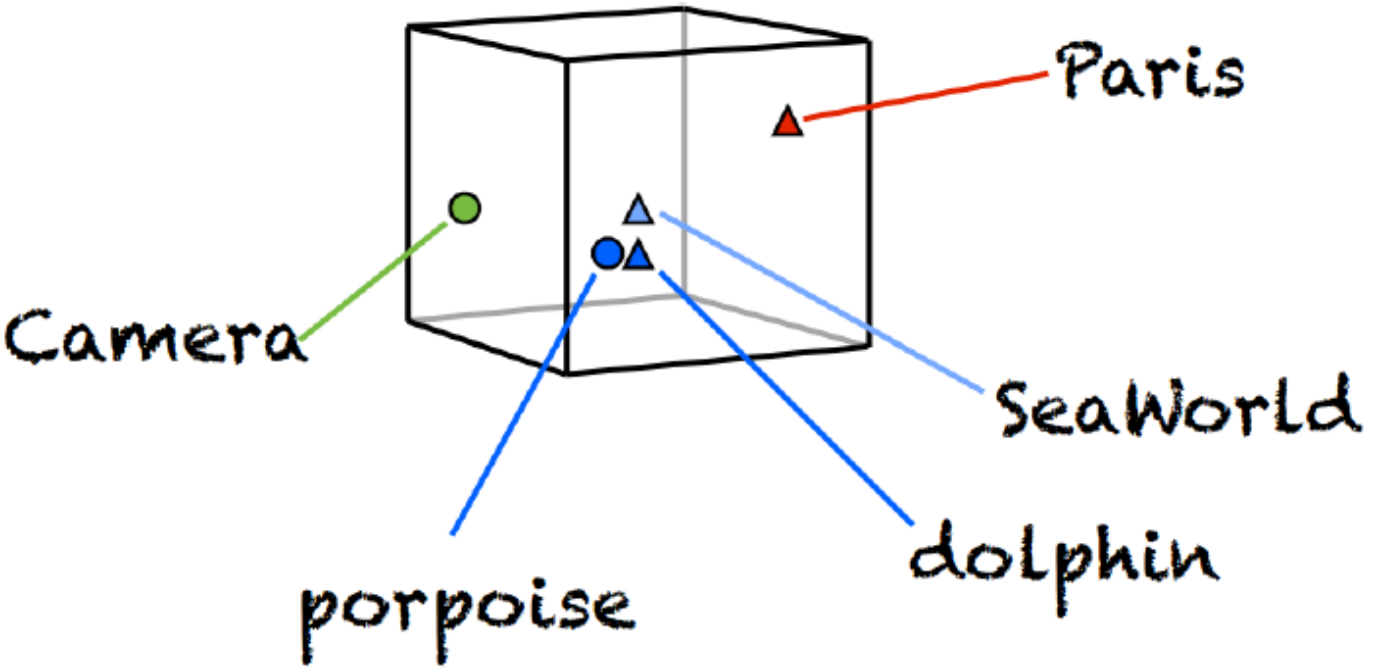
Embeddings are used to turn textual data (words, sentences, paragraphs) into high-dimensional vector representations and group them together with semantically similar data in a vectorspace. Thereby, computer can detect similarities mathematically.

Source

پردازش زبان طبیعی

جاسازی‌ها

NATURAL LANGUAGE PROCESSING – EMBEDDINGS



جاسازی‌ها، برای تبدیل داده‌های متنی (کلمات، جملات، پاراگراف‌ها) به بازنمایی‌های برداری با ابعاد بالا و گروه‌بندی آنها به داده‌های مشابه از نظر معنایی در یک فضای برداری استفاده می‌شوند.

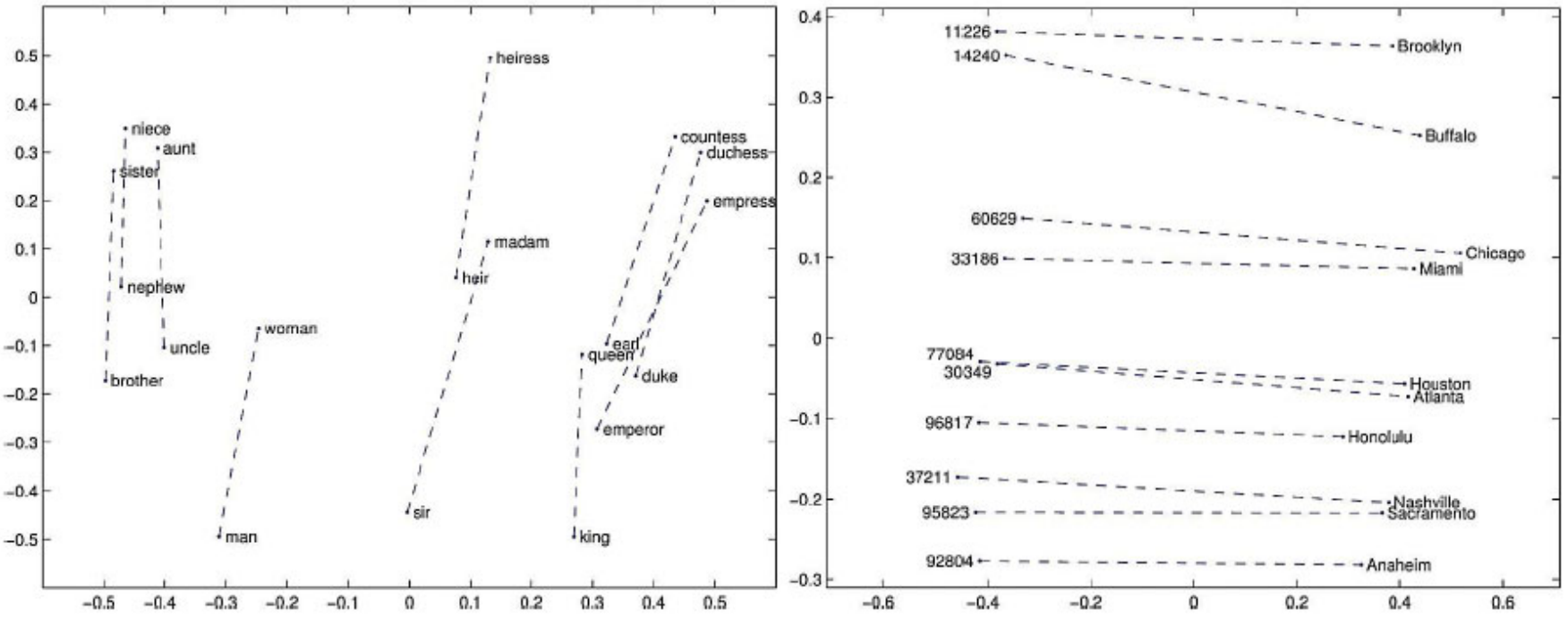
⇐ کامپیوتر می‌تواند شباهت‌ها را به‌صورت ریاضی آشکارسازی کند.

https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf

Deep Learning - Basics

Natural Language Processing – Word2Vec

Word2Vec is an unsupervised learning algorithm for obtaining vector representations for words. These vectors were trained for a specific domain on a very large textual data set. **GloVe** is a better performing alternative.



It detects similarities mathematically by grouping the vectors of similar words together. All it needs is words co-occurance in the given corpus.

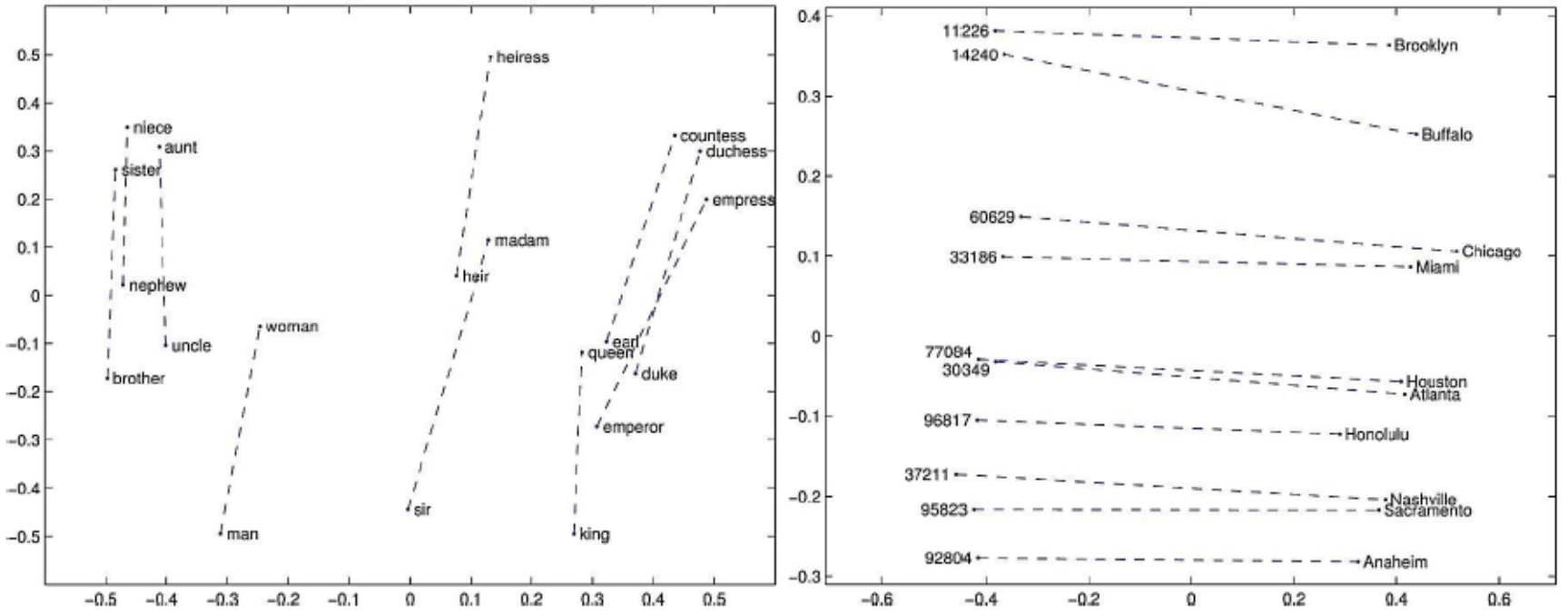
Source

پردازش زبان طبیعی

Word2Vec

NATURAL LANGUAGE PROCESSING – WORD2VEC

Word2Vec یک الگوریتم یادگیری بدون نظارت برای به‌دست آوردن بازنمایی برداری برای کلمات است.
این بردارها برای یک دامنه‌ی خاص بر روی یک مجموعه داده‌ی متنی بسیار بزرگ آموزش می‌بینند.
Glove یک جایگزین با کارآیی بهتر است.



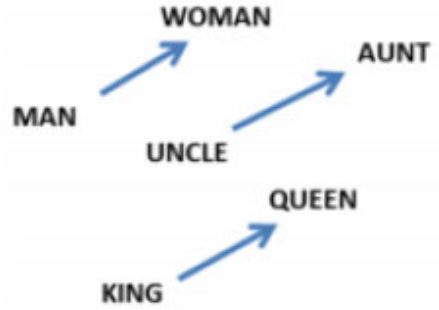
Word2Vec شباهت‌ها را به‌صورت ریاضی، با گروه‌بندی بردارهای کلمات مشابه با یکدیگر، تشخیص می‌دهد.
تمام آنچه این الگوریتم نیاز دارد، هم وقوعی کلمات در یک پیکره‌ی داده شده است.

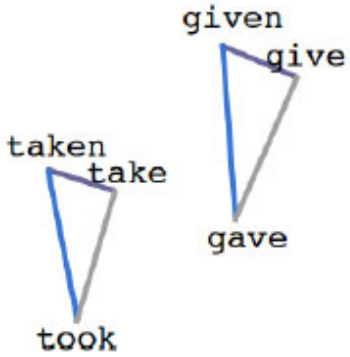
https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf

Deep Learning - Basics

Natural Language Processing – Word2Vec

FRANCE	JESUS	XBOX	REDDISH	SCRATCHED	MEGABITS
AUSTRIA	GOD	AMIGA	GREENISH	NAILED	OCTETS
BELGIUM	SATI	PLAYSTATION	BLUISH	SMASHED	MB/S
GERMANY	CHRIST	MSX	PINKISH	PUNCHED	BIT/S
ITALY	SATAN	IPOD	PURPLISH	POPPED	BAUD
GREECE	KALI	SEGA	BROWNISH	CRIMPED	CARATS
SWEDEN	INDRA	PSNUMBER	GREYISH	SCRAPED	KBIT/S
NORWAY	VISHNU	HD	GRAYISH	SCREWED	MEGAHERTZ
EUROPE	ANANDA	DREAMCAST	WHITISH	SECTIONED	MEGAPIXELS
HUNGARY	PARVATI	GEFORCE	SILVERY	SLASHED	GBIT/S
SWITZERLAND	GRACE	CAPCOM	YELLOWISH	RIPPED	AMPERES





Woman – Man ≈ Aunt - *Uncle*
King - Male + Female ≈ *Queen*
Human - Animal ≈ *Ethics*

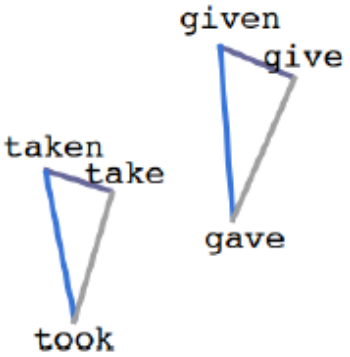
Source

پردازش زبان طبیعی

مثال : Word2Vec

NATURAL LANGUAGE PROCESSING – WORD2VEC

FRANCE	JESUS	XBOX	REDDISH	SCRATCHED	MEGABITS
AUSTRIA	GOD	AMIGA	GREENISH	NAILED	OCTETS
BELGIUM	SATI	PLAYSTATION	BLUISH	SMASHED	MB/S
GERMANY	CHRIST	MSX	PINKISH	PUNCHED	BIT/S
ITALY	SATAN	IPOD	PURPLISH	POPPED	BAUD
GREECE	KALI	SEGA	BROWNISH	CRIMPED	CARATS
SWEDEN	INDRA	PSNUMBER	GREYISH	SCRAPED	KBIT/S
NORWAY	VISHNU	HD	GRAYISH	SCREWED	MEGAHERTZ
EUROPE	ANANDA	DREAMCAST	WHITISH	SECTIONED	MEGAPIXELS
HUNGARY	PARVATI	GEFORCE	SILVERY	SLASHED	GBIT/S
SWITZERLAND	GRACE	CAPCOM	YELLOWISH	RIPPED	AMPERES



Woman – Man ≈ Aunt - *Uncle*
King - Male + Female ≈ *Queen*
Human - Animal ≈ *Ethics*

Deep Learning - Basics

Natural Language Processing – Thought Vectors

Thought vectors is a way of embedding thoughts in vector space. Their features will represent how each thought relates to other thoughts.

Convert every sentence in a document to a though vector, in a way that similar thoughts are nearby. You can do basic natural reasoning by learning to predict next thought vector based on a sequence of previous thought vectors. Thereby, by reading every document on the web, computers might be able to reason like humans do by mimicking the thoughts expressed in content.

A neural machine translation is trained on bilingual text using a encoder and decoder RNN. For translation, the input sentence is transformed into a thought vector. This vector is used to reconstruct the given thought in another language.

پردازش زبان طبیعی

بردارهای فکر

NATURAL LANGUAGE PROCESSING – THOUGHT VECTORS

بردارهای فکر، راهی برای جاسازی افکار در فضای برداری است.
ویژگی‌های آنها نشان می‌دهد چگونه هر فکر به سایر فکرها مرتبط می‌شود.

هر جمله در یک سند، به یک بردار فکر تبدیل می‌شود،
به‌شیوه‌ای که فکرهای مشابه، نزدیک هم باشند.
می‌توانیم یک استدلال طبیعی پایه را
از طریق یادگیری بردار فکر بعدی بر اساس یک دنباله از بردارهای فکر قبلی انجام بدهیم.
در نتیجه، با خواندن هر سند بر روی وب، کامپیوترها ممکن است بتوانند مانند انسان‌ها استدلال کنند
(با تقلید از فکرهای بیان شده در محتوای سندها).

یک ماشین ترجمه‌ی عصبی، با استفاده از یک شبکه‌ی عصبی بازگشتی کدگذار و گدگشا،
بر روی متن دوزبانه آموزش داده می‌شود.

برای ترجمه، جمله‌ی ورودی به یک بردار فکر تبدیل می‌شود.
این بردار فکر برای بازسازی فکر داده شده در زبان دیگر استفاده می‌شود.

مبانی یادگیری عمیق

۳

ملزومات یادگیری عمیق

Deep Learning - Basics

Usage Requirements

Large data set with good quality *(input-output mappings)*

Measurable and describable goals *(define the cost)*

Enough computing power *(AWS GPU Instance)*

Excels in tasks where the basic unit (pixel, word) has very little meaning in itself, but the combination of such units has a useful meaning.

ملزومات استفاده از یادگیری عمیق

DEEP LEARNING USAGE REQUIREMENT

وجود مجموعه داده‌ی بزرگ با کیفیت خوب
(نگاشت‌های ورودی–خروجی)

وجود هدف‌های قابل اندازه‌گیری و قابل توصیف
(برای تعریف تابع هزینه)

توان محاسباتی کافی
(برای نمونه: واحد پردازش گرافیکی / GPU)

یادگیری عمیق در وظیفه‌هایی برتری دارد که در آنها واحد پایه (پیکسل‌ها/کلمات)
به‌خودی‌خود معنای بسیار اندکی دارند،
اما ترکیب چنین واحدهایی دارای معنای مفیدی است.

مبانی یادگیری عمیق

۴

منابع

منبع اصلی

Deep Learning
The Past, Present and Future of
Artificial Intelligence
Lukas Masuch
@lukasmasuch
+lukasmasuch