

Data Mining

8 Chapter 8 Neural network and deep learning

Asso.Prof.Dr. Xiao-dong Zhu

Business School, University of Shanghai for Science & Technology

July, Summer 2018, Shanghai

Contents of chapter 8

- 1 Artificial Intelligence and Big Data Era
- 2 Artificial Neural Network
- 3 deep learning

Big data

We are now in a big data era.

- large scale data processing: storage and computing
- mobile internet. Everybody makes data.
- social network. Facebook, twitter, wechat, weibo.



For instance, in big data era

- 6 billion new photos are reported every month by Facebook.

For instance, in big data era

- 6 billion new photos are reported every month by Facebook.
- 72 hours of video are uploaded to YouTube every minute. (in 2012, 48 hours in 2011)

For instance, in big data era

- 6 billion new photos are reported every month by Facebook.
- 72 hours of video are uploaded to YouTube every minute. (in 2012, 48 hours in 2011)
- 204 million email message are send. (in 2012)

For instance, in big data era

- 6 billion new photos are reported every month by Facebook.
- 72 hours of video are uploaded to YouTube every minute. (in 2012, 48 hours in 2011)
- 204 million email message are send. (in 2012)
- ...

For instance, in big data era

- 6 billion new photos are reported every month by Facebook.
- 72 hours of video are uploaded to YouTube every minute. (in 2012, 48 hours in 2011)
- 204 million email message are send. (in 2012)
- ...

For instance, in big data era

- 6 billion new photos are reported every month by Facebook.
- 72 hours of video are uploaded to YouTube every minute. (in 2012, 48 hours in 2011)
- 204 million email message are send. (in 2012)
- ...



prediction

- According to IDC, in 2011 we created 1.8 zettabytes (or 1.8 trillion GBs) of information, which is enough data to fill 57.5 billion 32GB Apple iPads. That's enough iPads to build a Great iPad Wall of China twice as tall as the original.
- In 2012 it reached 2.8 zettabytes
- IDC now forecasts that we will generate 40 zettabytes (ZB) by 2020.

prediction

- According to IDC, in 2011 we created 1.8 zettabytes (or 1.8 trillion GBs) of information, which is enough data to fill 57.5 billion 32GB Apple iPads. That's enough iPads to build a Great iPad Wall of China twice as tall as the original.
- In 2012 it reached 2.8 zettabytes
- IDC now forecasts that we will generate 40 zettabytes (ZB) by 2020.

IDC, International Data Corporation, is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets.

Big data has 4 characters

Volume

The quantity of generated data is huge

Big data has 4 characters

Volume

The quantity of generated data is huge

Velocity

The speed of generating data is so fast

Big data has 4 characters

Volume

The quantity of generated data is huge

Velocity

The speed of generating data is so fast

Variety

text, image, audio, video, ...

Big data has 4 characters

Volume

The quantity of generated data is huge

Velocity

The speed of generating data is so fast

Variety

text, image, audio, video, ...

Veracity

The data is fuzzy, uncertain

Big data has 4 characters

Volume

The quantity of generated data is huge

Velocity

The speed of generating data is so fast

Variety

text, image, audio, video, ...

Veracity

The data is fuzzy, uncertain

Value

After mining, expect valuable experience, knowledge, patterns, rules...

volume

- 1 Byte = 8 bit

volume

- 1 Byte = 8 bit
- 1 KB = 2^{10} Bytes = 1,024 Bytes

volume

- 1 Byte = 8 bit
- 1 KB = 2^{10} Bytes = 1,024 Bytes
- 1 MB = 2^{10} KB = 2^{20} Bytes = 1,048,576 Bytes

volume

- 1 Byte = 8 bit
- 1 KB = 2^{10} Bytes = 1,024 Bytes
- 1 MB = 2^{10} KB = 2^{20} Bytes = 1,048,576 Bytes
- 1 GB = 2^{10} MB = 2^{30} Bytes = 1,073,741,824 Bytes

volume

- 1 Byte = 8 bit
- 1 KB = 2^{10} Bytes = 1,024 Bytes
- 1 MB = 2^{10} KB = 2^{20} Bytes = 1,048,576 Bytes
- 1 GB = 2^{10} MB = 2^{30} Bytes = 1,073,741,824 Bytes
- 1 TB = 2^{10} GB = 2^{40} Bytes = 1,099,511,627,776 Bytes

volume

- 1 Byte = 8 bit
- 1 KB = 2^{10} Bytes = 1,024 Bytes
- 1 MB = 2^{10} KB = 2^{20} Bytes = 1,048,576 Bytes
- 1 GB = 2^{10} MB = 2^{30} Bytes = 1,073,741,824 Bytes
- 1 TB = 2^{10} GB = 2^{40} Bytes = 1,099,511,627,776 Bytes
- 1 PB = 2^{10} TB = 2^{50} Bytes = 1,125,899,906,842,624 Bytes

volume

- 1 Byte = 8 bit
- 1 KB = 2^{10} Bytes = 1,024 Bytes
- 1 MB = 2^{10} KB = 2^{20} Bytes = 1,048,576 Bytes
- 1 GB = 2^{10} MB = 2^{30} Bytes = 1,073,741,824 Bytes
- 1 TB = 2^{10} GB = 2^{40} Bytes = 1,099,511,627,776 Bytes
- 1 PB = 2^{10} TB = 2^{50} Bytes = 1,125,899,906,842,624 Bytes
- 1 EB = 2^{10} PB = 2^{60} Bytes = 1,152,921,504,606,846,976 Bytes

volume

- 1 Byte = 8 bit
- 1 KB = 2^{10} Bytes = 1,024 Bytes
- 1 MB = 2^{10} KB = 2^{20} Bytes = 1,048,576 Bytes
- 1 GB = 2^{10} MB = 2^{30} Bytes = 1,073,741,824 Bytes
- 1 TB = 2^{10} GB = 2^{40} Bytes = 1,099,511,627,776 Bytes
- 1 PB = 2^{10} TB = 2^{50} Bytes = 1,125,899,906,842,624 Bytes
- 1 EB = 2^{10} PB = 2^{60} Bytes = 1,152,921,504,606,846,976 Bytes
- 1 ZB = 2^{10} EB = 2^{70} Bytes = 1,180,591,620,717,411,303,424 Bytes

volume

- 1 Byte = 8 bit
- 1 KB = 2^{10} Bytes = 1,024 Bytes
- 1 MB = 2^{10} KB = 2^{20} Bytes = 1,048,576 Bytes
- 1 GB = 2^{10} MB = 2^{30} Bytes = 1,073,741,824 Bytes
- 1 TB = 2^{10} GB = 2^{40} Bytes = 1,099,511,627,776 Bytes
- 1 PB = 2^{10} TB = 2^{50} Bytes = 1,125,899,906,842,624 Bytes
- 1 EB = 2^{10} PB = 2^{60} Bytes = 1,152,921,504,606,846,976 Bytes
- 1 ZB = 2^{10} EB = 2^{70} Bytes =
1,180,591,620,717,411,303,424 Bytes
- 1 YB = 2^{10} ZB = 2^{80} Bytes =
1,208,925,819,614,629,174,706,176 Bytes

In Wikipedia

Artificial intelligence (AI), sometimes called machine intelligence, is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans and other animals.

Nils J. Nilsson-Stanford University

artificial intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment.

In computer science

AI research is defined as the study of "intelligent agents": any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals.

research scope of A.I.

- Natural language processing,
- knowledge representation,
- intelligent search,
- reasoning,
- planning,
- machine learning,
- knowledge acquisition,
- combined scheduling problem,

research scope of A.I.

- Natural language processing,
- knowledge representation,
- intelligent search,
- reasoning,
- planning,
- machine learning,
- knowledge acquisition,
- combined scheduling problem,
- perception problem,
- pattern recognition,
- logic programming soft computing,
- inaccurate and uncertain management,
- artificial life,
- neural network,
- Complex system,
- genetic algorithm

Neural Networks can be

- 1 **Biological** models
- 2 **Artificial** models

Desire to produce **artificial systems** capable of sophisticated computations similar to the human brain.

Biological analogy and some concepts

- the brain is composed of a mass of interconnected neurons
- neurons transmits signals to each other
- whether a signal is transmitted is an all-or-nothing event.(the electrical potential in the neuron cell body is thresholded)
- whether a signal is sent, depends on the strength of the bond(synapse) between two neurons.

How does the brain work?

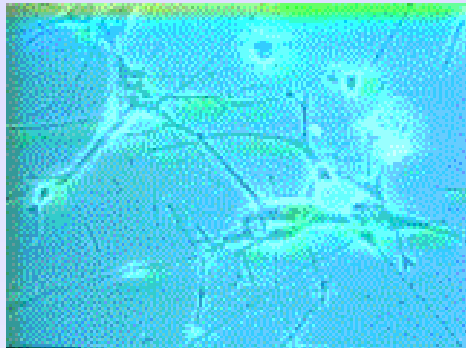
neuron

- the cell that performs information processing in the brain
- fundamental functional unit of all nervous system tissue

How does the brain work?

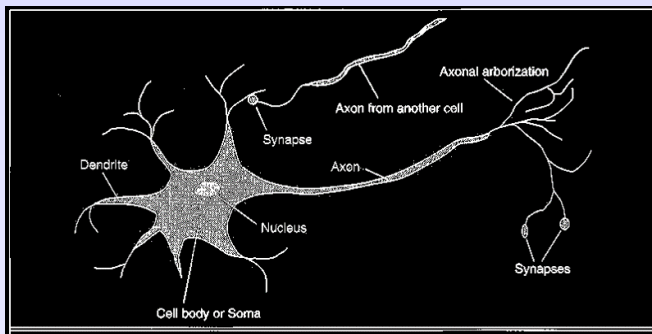
neuron

- the cell that performs information processing in the brain
- fundamental functional unit of all nervous system tissue



How does the brain work?

each neuron consists of: Soma(躯体), dendrites(树突), axon(轴突), axonal arborization(轴突的树枝状), synapse(突触)



In fact, the length of the axon should be about 100 times the diameter of the cell body. (from brain to toes)

Comparison of brain and computer

Table: Comparison of brain and computer

	Human	Computer
processing elements	100 billion neurons	10 million gates
interconnects	1000 per neuron	a few
cycles per sec	1000	500 million
evolution	200000 years	2 years

history

- **1943**, McCulloch & Pitts show that neurons can be combined to construct a Turing machine (using ANDs, Ors, & NOTs)

history

- 1943, McCulloch & Pitts show that neurons can be combined to construct a Turing machine (using ANDs, Ors, & NOTs)
- 1958, Rosenblatt shows that perceptrons will converge if what they are trying to learn can be represented

history

- 1943, McCulloch & Pitts show that neurons can be combined to construct a Turing machine (using ANDs, Ors, & NOTs)
- 1958, Rosenblatt shows that perceptrons will converge if what they are trying to learn can be represented
- 1969, Minsky & Papert showed the limitations of perceptrons, killing research for a decade

history

- 1943, McCulloch & Pitts show that neurons can be combined to construct a Turing machine (using ANDs, Ors, & NOTs)
- 1958, Rosenblatt shows that perceptrons will converge if what they are trying to learn can be represented
- 1969, Minsky & Papert showed the limitations of perceptrons, killing research for a decade
- 1985, backpropagation algorithm revitalizes the field

history

- 1943, McCulloch & Pitts show that neurons can be combined to construct a Turing machine (using ANDs, Ors, & NOTs)
- 1958, Rosenblatt shows that perceptrons will converge if what they are trying to learn can be represented
- 1969, Minsky & Papert showed the limitations of perceptrons, killing research for a decade
- 1985, backpropagation algorithm revitalizes the field
- 1980s-, machine learning, realizing artificial intelligence (speech recognition, image recognition, playing go, dialogue)

history

- **1943**, McCulloch & Pitts show that neurons can be combined to construct a Turing machine (using ANDs, Ors, & NOTs)
- **1958**, Rosenblatt shows that perceptrons will converge if what they are trying to learn can be represented
- **1969**, Minsky & Papert showed the limitations of perceptrons, killing research for a decade
- **1985**, backpropagation algorithm revitalizes the field
- **1980s-**, machine learning, realizing artificial intelligence (speech recognition, image recognition, playing go, dialogue)
- **mid 90's**, Cortes & Vapnik published paper on Support Vector Machines

history

- **1943**, McCulloch & Pitts show that neurons can be combined to construct a Turing machine (using ANDs, Ors, & NOTs)
- **1958**, Rosenblatt shows that perceptrons will converge if what they are trying to learn can be represented
- **1969**, Minsky & Papert showed the limitations of perceptrons, killing research for a decade
- **1985**, backpropagation algorithm revitalizes the field
- **1980s-**, machine learning, realizing artificial intelligence (speech recognition, image recognition, playing go, dialogue)
- **mid 90's**, Cortes & Vapnik published paper on Support Vector Machines
- **2006-**, deep learning, for machine learning for higher prediction accuracy

history

- 2007, Fei-Fei Li's ImageNet assembling a database of 14 million labeled images (Data drives learning)

history

- **2007**, Fei-Fei Li's ImageNet assembling a database of 14 million labeled images (Data drives learning)
- **2011**, Microsoft explored Speech recognition

history

- **2007**, Fei-Fei Li's ImageNet assembling a database of 14 million labeled images (Data drives learning)
- **2011**, Microsoft explored Speech recognition
- **2014**, Google acquired DeepMind, combining deep learning and reinforcement learning

history

- 2007, Fei-Fei Li's ImageNet assembling a database of 14 million labeled images (Data drives learning)
- 2011, Microsoft explored Speech recognition
- 2014, Google acquired DeepMind, combining deep learning and reinforcement learning
- 2016, DeepMind's AlphaGo defeated world champion Lee Sedol by 4-1

history

- 2007, Fei-Fei Li's ImageNet assembling a database of 14 million labeled images (Data drives learning)
- 2011, Microsoft explored Speech recognition
- 2014, Google acquired DeepMind, combining deep learning and reinforcement learning
- 2016, DeepMind's AlphaGo defeated world champion Lee Sedol by 4-1
- 2017, DeepMind's AlphaGo won Chinese Go champion Ke Jie by 3-0

definition of artificial neural network

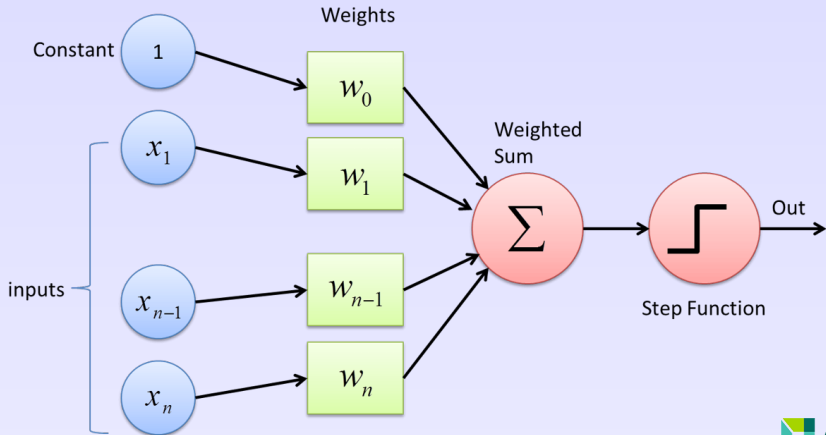
A Artificial Neural Network is a system composed of many simple processing elements operating in parallel which can acquire, store, and utilize experiential knowledge.

definition of artificial neural network

A Artificial Neural Network is a system composed of many simple processing elements operating in parallel which can acquire, store, and utilize experiential knowledge.

- Each element of NN is a node called unit.
- Units are connected by links.
- Each link has a numeric weight.

a N.N. with 4-layer, several input units, and 2 output units



TRAINING

During the training phase, a neural network is fed thousands of labeled images of various animals, learning to classify them.



INPUT

An unlabeled image is shown to the pretrained network.



FIRST LAYER

The neurons respond to different simple shapes, like edges.



HIGHER LAYER

Neurons respond to more complex structures.



TOP LAYER

Neurons respond to highly complex, abstract concepts that we would identify as different animals.



OUTPUT

The network predicts what the object most likely is, based on its training.

90% DOG



10% WOLF

drive behind of deep learning

- faster computing devices and cores (CPU/GPU)
- larger storage and memory.
- big data.
- cloud computing.
- new models and algorithms.

cloud

Companies

Google, Amazon, Facebook, Baidu, Tencent, Alibaba, et al.



cloud

Companies

Google, Amazon, Facebook, Baidu, Tencent, Alibaba, et al.

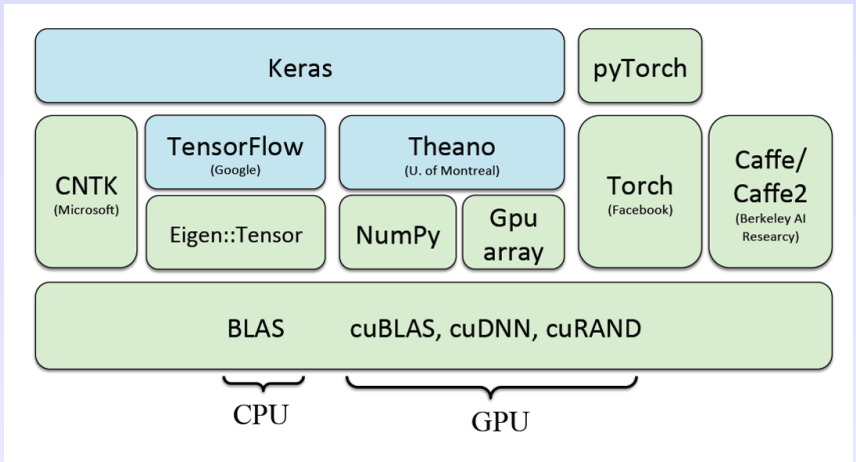
Cloud for ...

- 1 cloud computing
- 2 cloud storage

How to implement deep learning

- linux cluster with GPU servers
- Python language
- TensorFlow, Keras, Theano (packages)

deep learning software



application of deep learning

- Information retrieval (search engines)

application of deep learning

- Information retrieval (search engines)
- Pattern recognition

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)



application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playing

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playingp
- Natural Language Processing

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playingp
- Natural Language Processing
- Social media mining

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playing
- Natural Language Processing
- Social media mining
- robot customer service

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playing
- Natural Language Processing
- Social media mining
- robot customer service

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playingp
- Natural Language Processing
- Social media mining
- robot customer service
- Brand and product differentiation

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playing
- Natural Language Processing
- Social media mining
- robot customer service
- Brand and product differentiation
- Language Translation

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playing
- Natural Language Processing
- Social media mining
- robot customer service
- Brand and product differentiation
- Language Translation
- Speech Recognition

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playing
- Natural Language Processing
- Social media mining
- robot customer service
- Brand and product differentiation
- Language Translation
- Speech Recognition
- Generating Handwriting

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playing
- Natural Language Processing
- Social media mining
- robot customer service
- Brand and product differentiation
- Language Translation
- Speech Recognition
- Generating Handwriting
- Face Recognition

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playingp
- Natural Language Processing
- Social media mining
- robot customer service
- Brand and product differentiation
- Language Translation
- Speech Recognition
- Generating Handwriting
- Face Recognition
- Autonomous Driving

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playingp
- Natural Language Processing
- Social media mining
- robot customer service
- Brand and product differentiation
- Language Translation
- Speech Recognition
- Generating Handwriting
- Face Recognition
- Autonomous Driving
- Generating Arts

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playingp
- Natural Language Processing
- Social media mining
- robot customer service
- Brand and product differentiation
- Language Translation
- Speech Recognition
- Generating Handwriting
- Face Recognition
- Autonomous Driving
- Generating Arts
- Imitating Famous Painters

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playingp
- Natural Language Processing
- Social media mining
- robot customer service
- Brand and product differentiation
- Language Translation
- Speech Recognition
- Generating Handwriting
- Face Recognition
- Autonomous Driving
- Generating Arts
- Imitating Famous Painters
- Generating Music

application of deep learning

- Information retrieval (search engines)
- Pattern recognition
- customer targeting
- Sentiment analysis (based on written text)
- Personalization recommendation
- Chess playingp
- Natural Language Processing
- Social media mining
- robot customer service
- Brand and product differentiation
- Language Translation
- Speech Recognition
- Generating Handwriting
- Face Recognition
- Autonomous Driving
- Generating Arts
- Imitating Famous Painters
- Generating Music
- Generating Photos

in Power system

- Peak Load Forecast (e.g. Maximum Demand)
- Failure Prediction (e.g. Battery)
- Condition Monitoring (e.g. Partial Discharge)