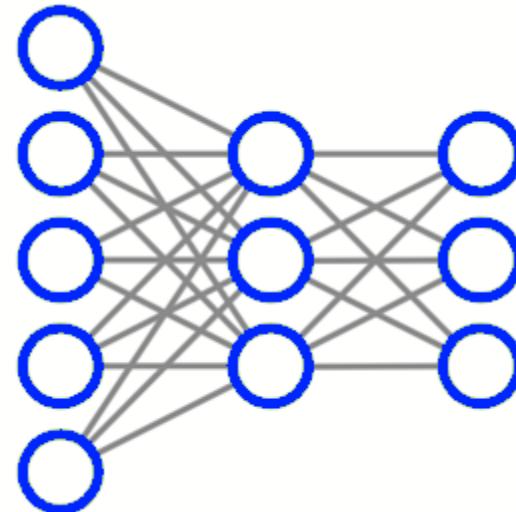


Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ciencia de la Computación



IIC1005 – Computación: Ciencia y Tecnología del Mundo Digital

Introducción a Deep Learning



Hans Löbel

¿Why a class solely focused on Deep Learning?

(aka why is Deep Learning so important?)

AlphaGo seals 4-1 victory over Go grandmaster Lee Sedol

DeepMind's artificial intelligence astonishes fans to defeat human opponent and offers evidence computer software has mastered a major challenge



Driverless pods plot new course to overtake humans

Autonomous cars used at Heathrow and being trialled in south-east London now beg the question - should humans be banned from driving?

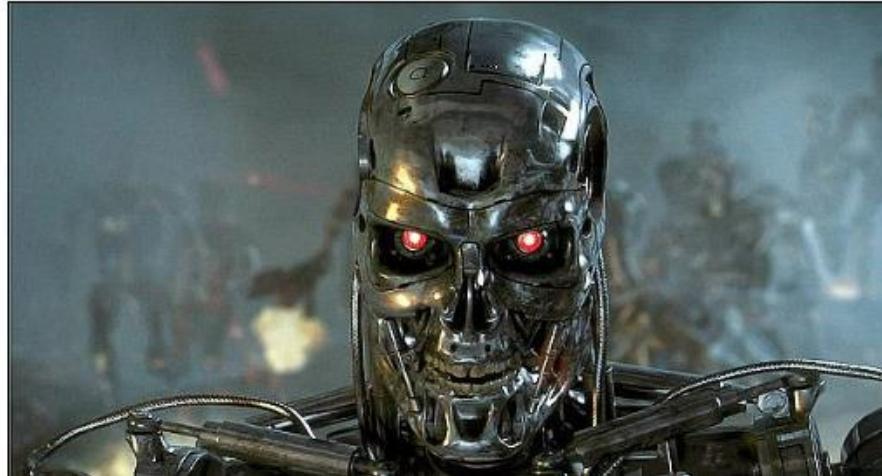


¿Why a class solely focused on Deep Learning?

(aka why is Deep Learning so important?)

Rise of the machines: Google AI experiment may lead to robots that can learn WITHOUT human input

- Generative Adversarial Networks create digital content based on real-life
- Google project pits AI algorithms against each other to refine this output
- The results could one day lead to machines that can learn without human input



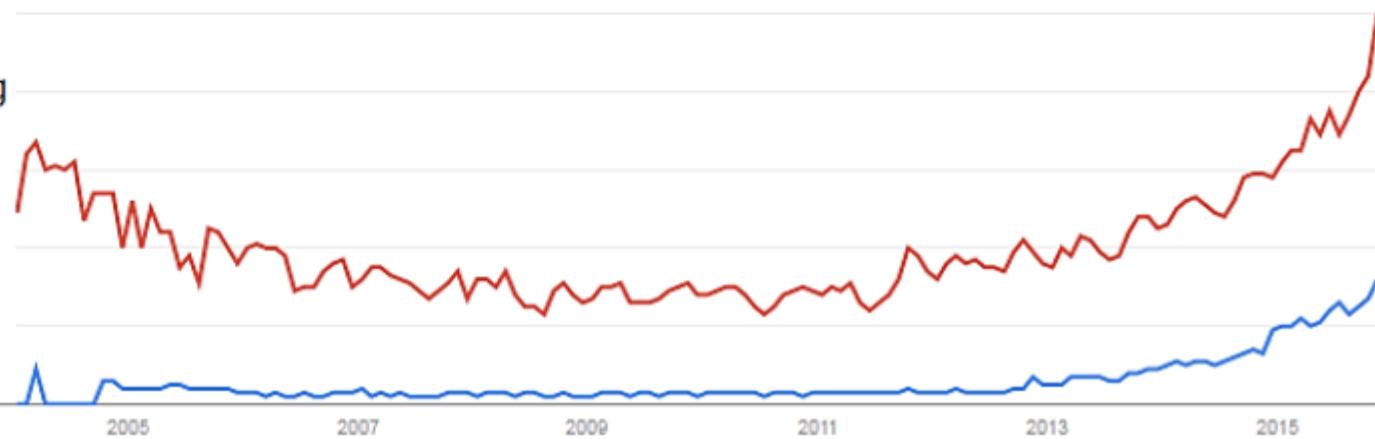


deep learning

Search term

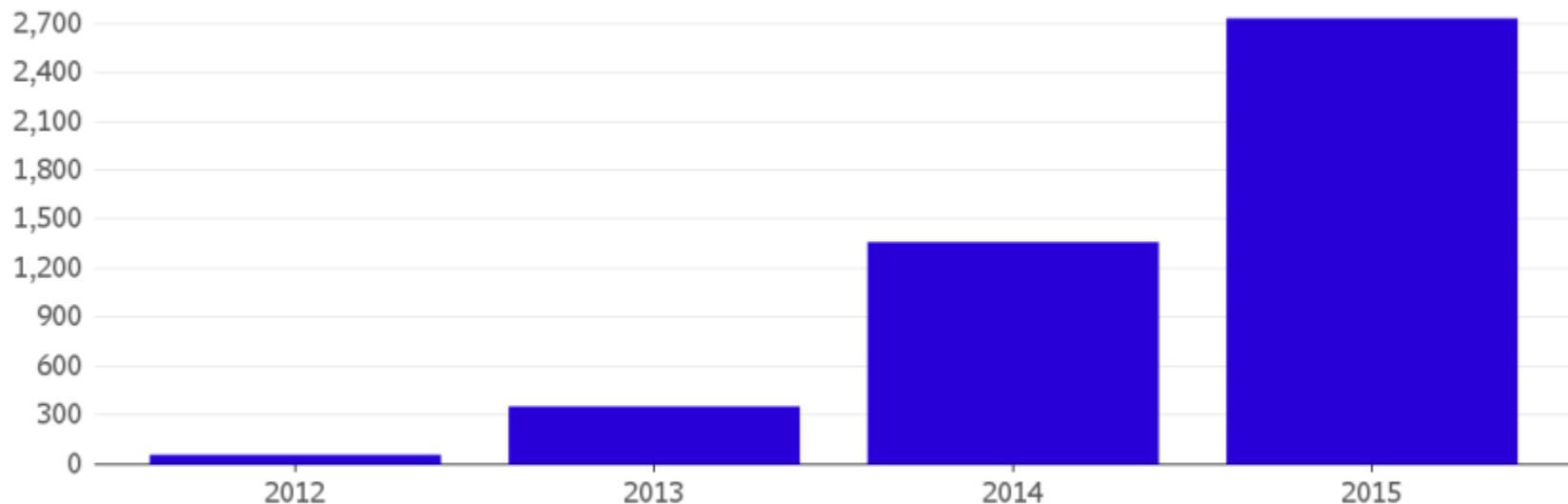
machine learning

Search term



Artificial Intelligence Takes Off at Google

Number of software projects within Google that uses a key AI technology, called Deep Learning.



Source: Google

Note: 2015 data does not incorporate data from Q4



Andrej Karpathy ✅

@karpathy

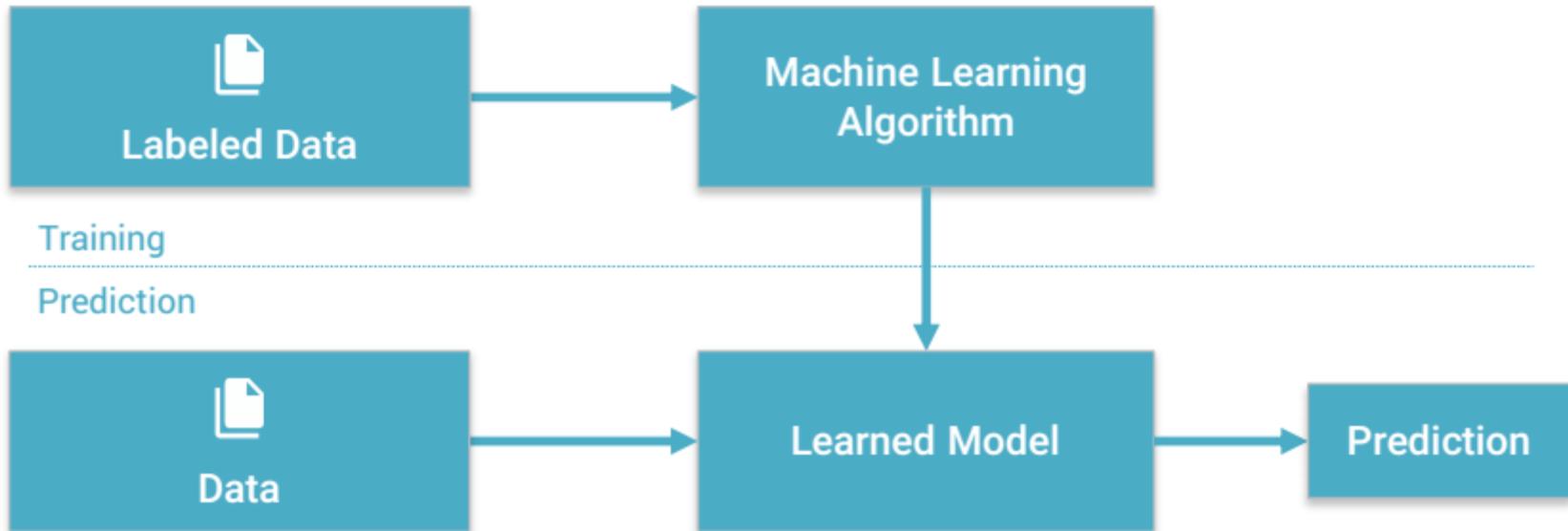
Following

Came to visit first class of [@cs231n](#) at Stanford. 2015: 150 students, 2016: 350, this year: 750. [#aiinterestsingularity](#)





Machine Learning is a type of Artificial Intelligence that provides computers with the ability to **learn without being explicitly programmed**.



Provides **various techniques** that can learn from and make predictions on data

Deep Learning is currently the most popular tool/technique in Machine Learning



Part of the machine learning field of learning representations of data. Exceptional effective at learning patterns.

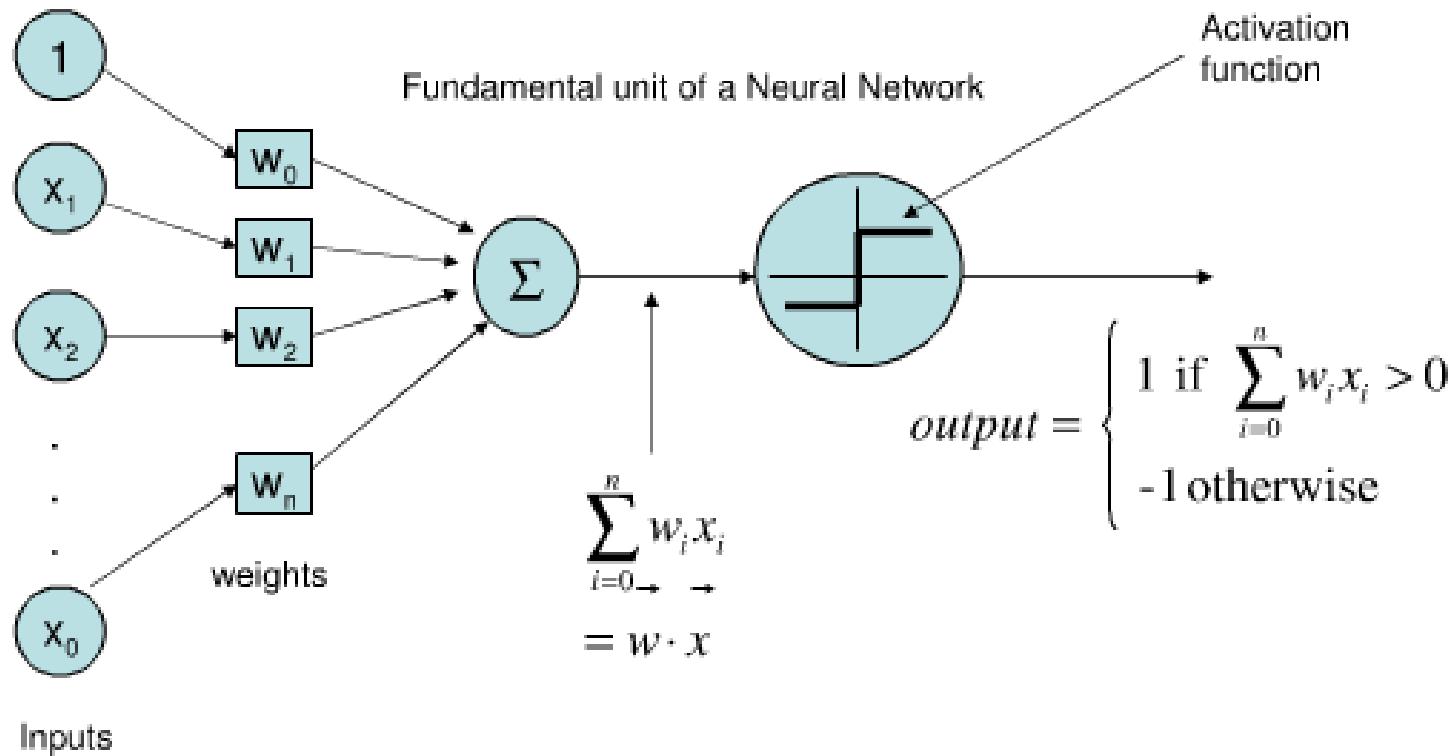


Utilizes learning algorithms that derive meaning out of data by using a **hierarchy** of multiple layers that **mimic the neural networks of our brain**.

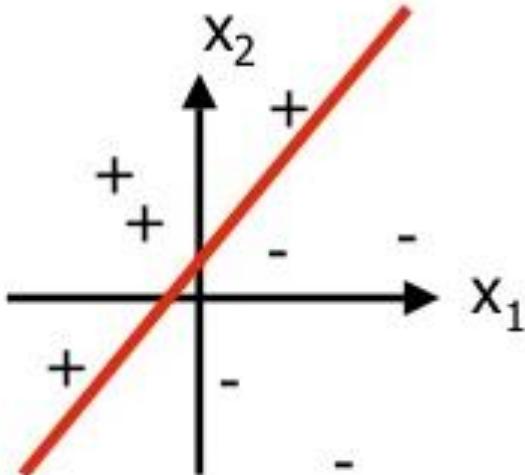


If you provide the system tons of information, it begins to understand it and respond in useful ways.

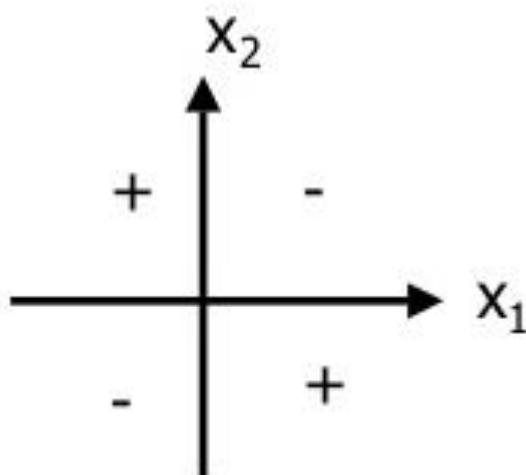
All started with a tiny Perceptron



...but it wasn't very powerful



Linearly separable



Non-Linearly separable



1958 Perceptron

1974 Backpropagation



Convolution Neural Networks for
Handwritten Recognition



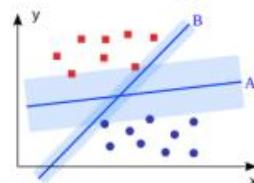
Google Brain Project on
16k Cores
2012



1998

awkward silence (AI Winter)

1995
SVM reigns



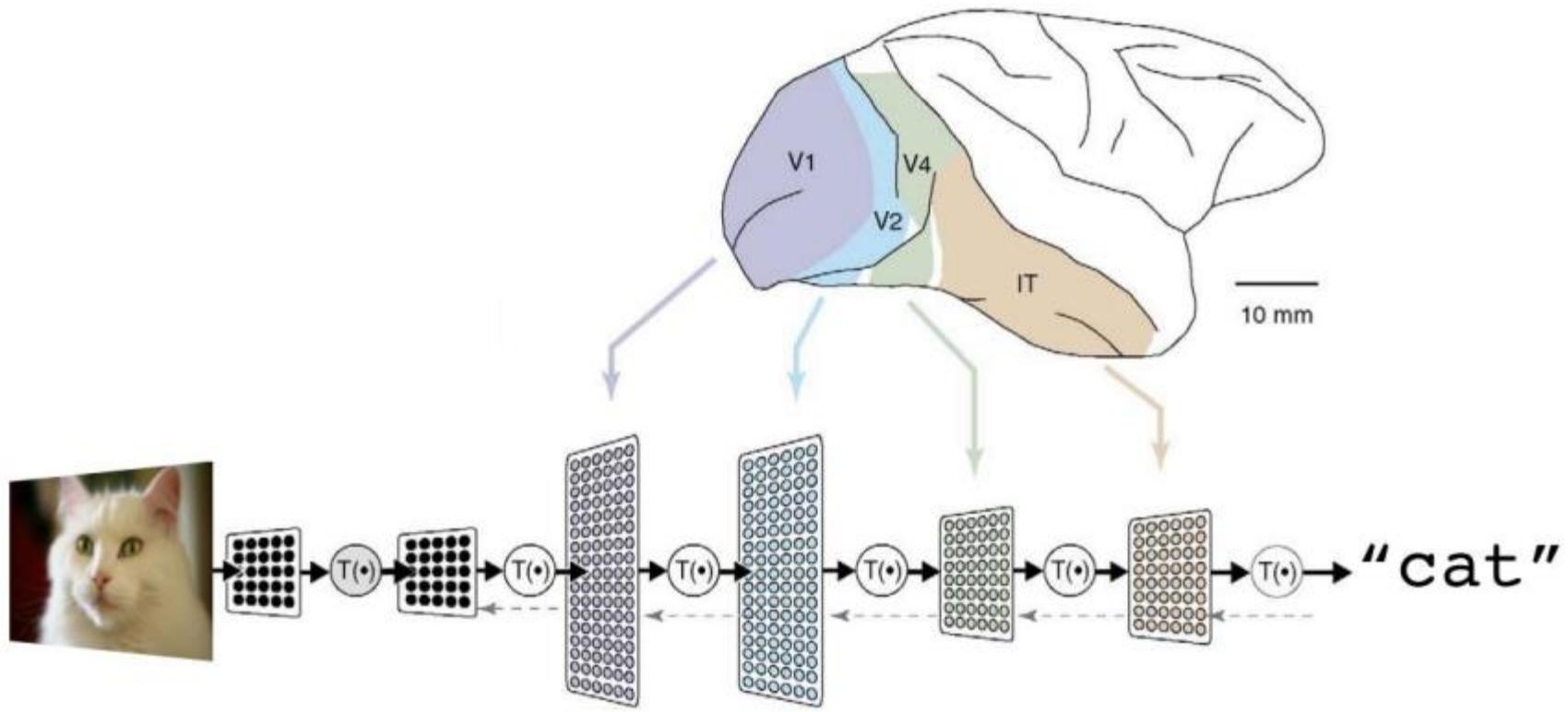
2006
Restricted
Boltzmann
Machine



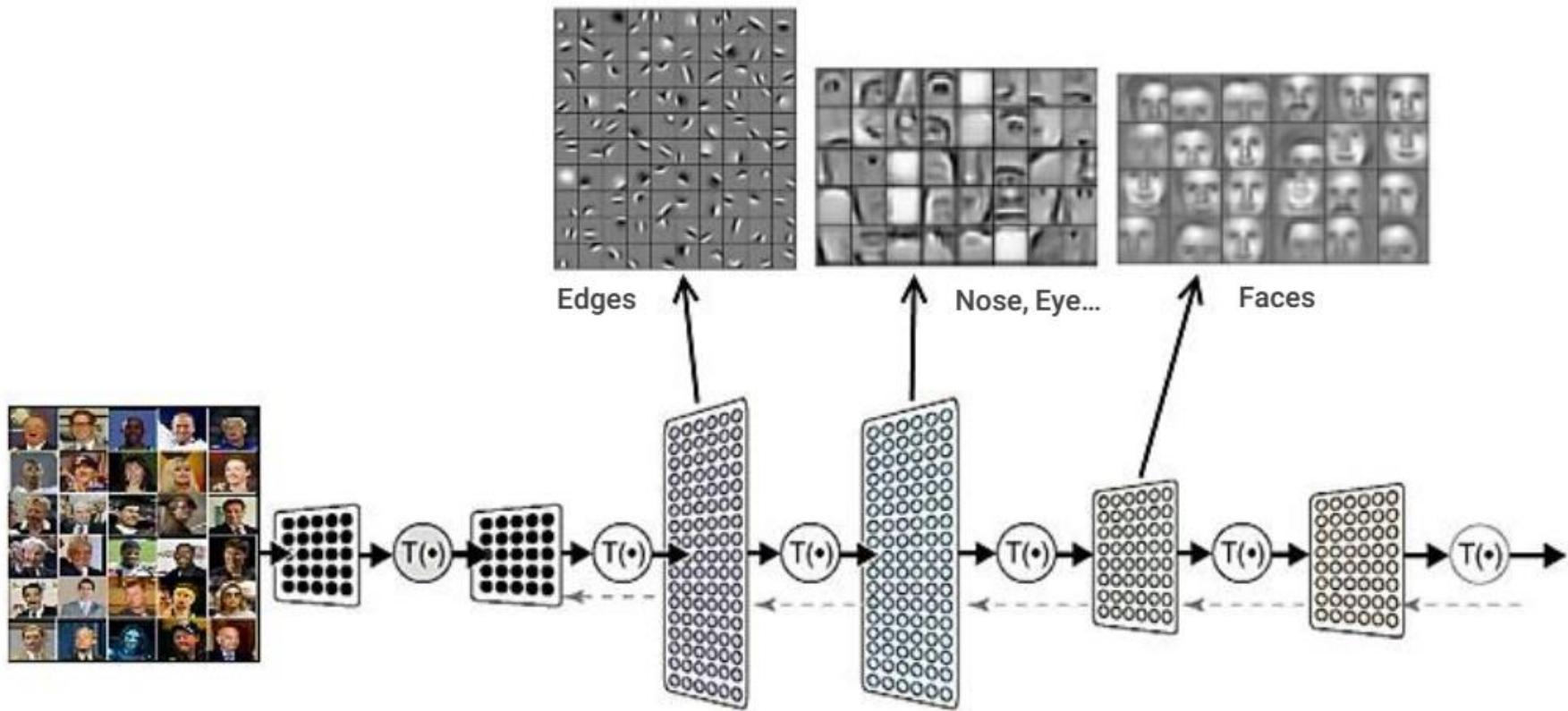
2012
AlexNet wins
ImageNet
IMAGENET

1969
Perceptron criticized

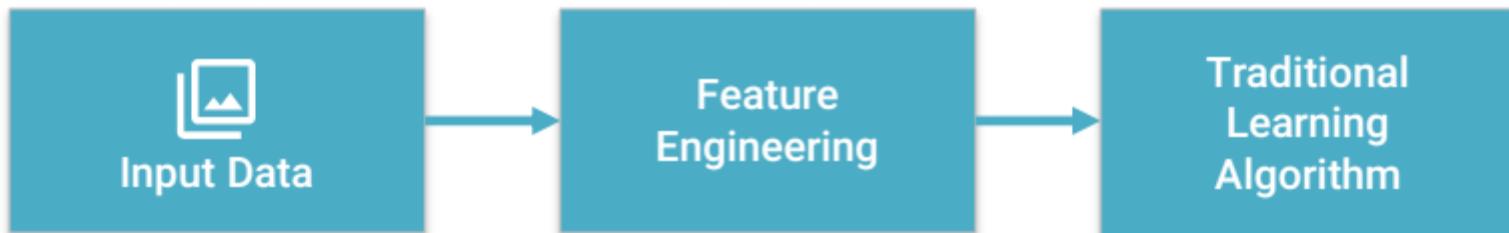




A deep neural network consists of a **hierarchy of layers**, whereby each layer **transforms the input data** into more abstract representations (e.g. edge -> nose -> face). The output layer combines those features to make predictions.



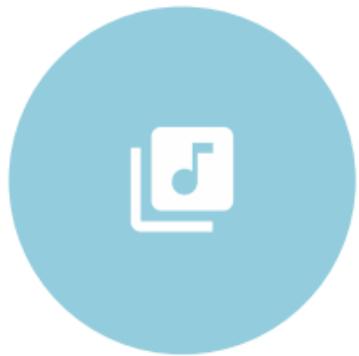
Deep Learning's power comes from its ability to learn a hierarchy of meaningful representations from data



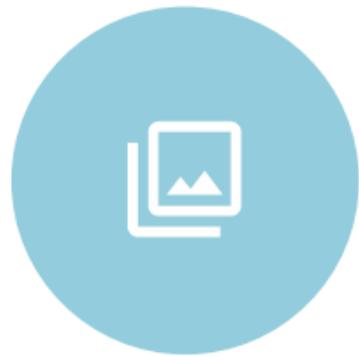
Costs lots of time



Unbeatable in many domains



Speech
Recognition

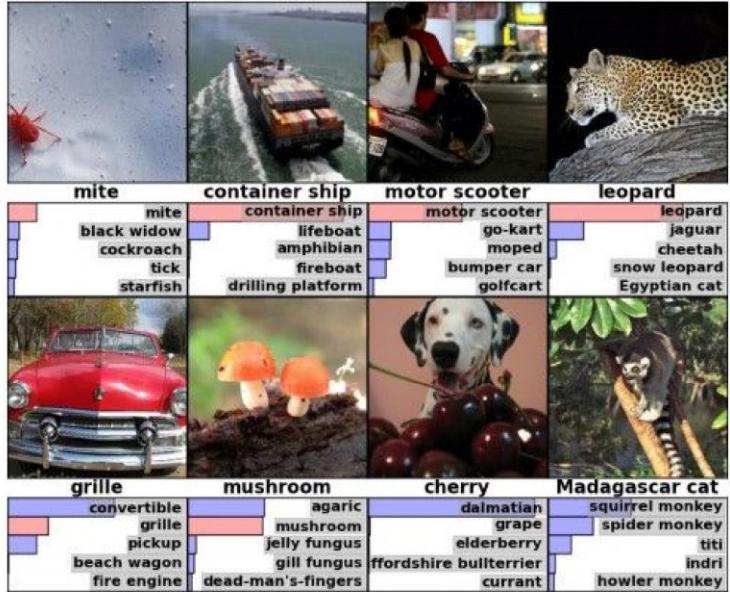


Computer
Vision



Natural Language
Processing

Classification

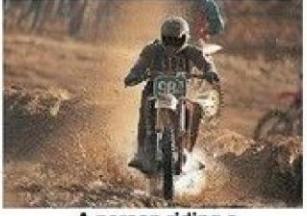


Retrieval



[Krizhevsky 2012]

Image Captioning

| Describes without errors | Describes with minor errors | Somewhat related to the image | Unrelated to the image |
|--|--|---|--|
|  |  |  |  |
| A person riding a motorcycle on a dirt road. | Two dogs play in the grass. | A skateboarder does a trick on a ramp. | A dog is jumping to catch a frisbee. |
|  |  |  |  |
| A group of young people playing a game of frisbee. | Two hockey players are fighting over the puck. | A little girl in a pink hat is blowing bubbles. | A refrigerator filled with lots of food and drinks. |
|  |  |  |  |
| A herd of elephants walking across a dry grass field. | A close up of a cat laying on a couch. | A red motorcycle parked on the side of the road. | A yellow school bus parked in a parking lot. |

[Vinyals et al., 2015]

Query Image



Retrieved Recipe

| Ingredients | Instructions |
|----------------|---|
| pasta | 1. Preheat oven to 350F. |
| ground beef | 2. Boil pasta until just cooked. |
| taco seasoning | 3. Brown ground beef and then drain. |
| water | 4. Add taco seasoning and water to meat and simmer for 5 minutes. |
| cream cheese | ... |
| cheese | 5. Put half of the shredded cheese over pasta, then cover with hamburger meat and mix gentle. |
| | 6. Sprinkle remaining cheese over the top. |
| | 7. Cook in the oven uncovered for 15-20 minutes. |

Query Recipe

| Ingredients | Instructions |
|-------------------|---|
| hamburger | 1. Cook hamburger until done and drain off the fat. |
| rigatoni pasta | 2. Add mushrooms and onion and fry until translucent. |
| Ragu pizza sauce | 3. Add pepperoni. |
| mushrooms | 4. ... |
| onion | 5. Lay noodles on top of hamburger mix in crockpot. |
| pepperoni | 6. Turn crock on low and leave 4-5 hours. |
| mozzarella cheese | 7. Pour over the remainder of pizza sauce over the noodles. |
| | 8. Top with the cheese. |



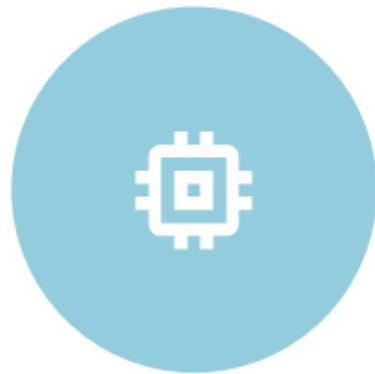
Retrieved Image



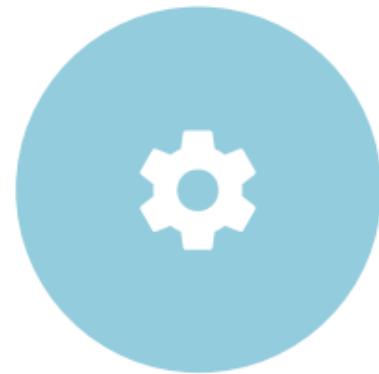
The right elements at the right time



Big Data
(Digitalization)



Computation
(Moore's Law, GPUs)



Algorithmic
Progress

There won't be a new “AI winter” (hopefully)

facebook



YAHOO!

Google



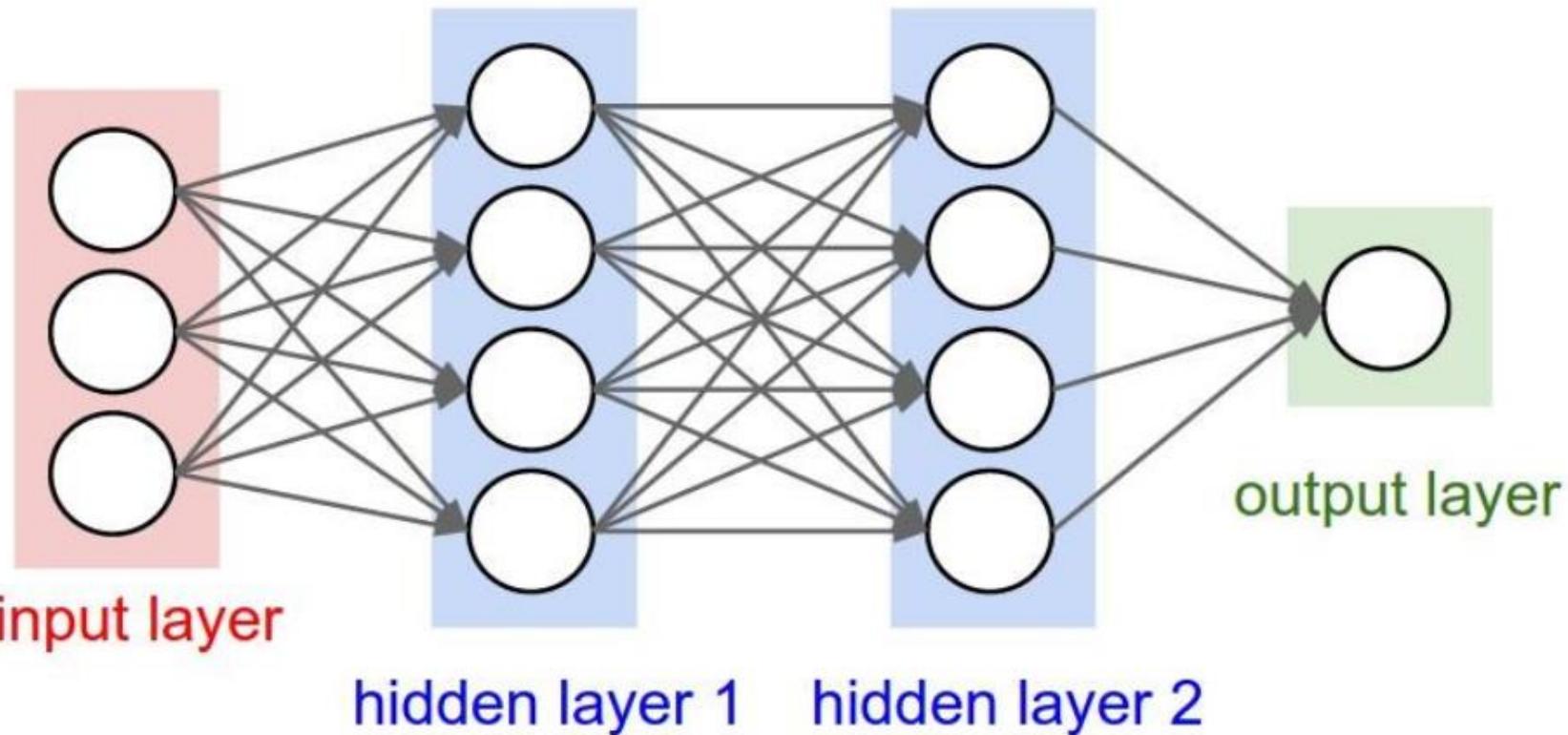
IBM



NVIDIA.

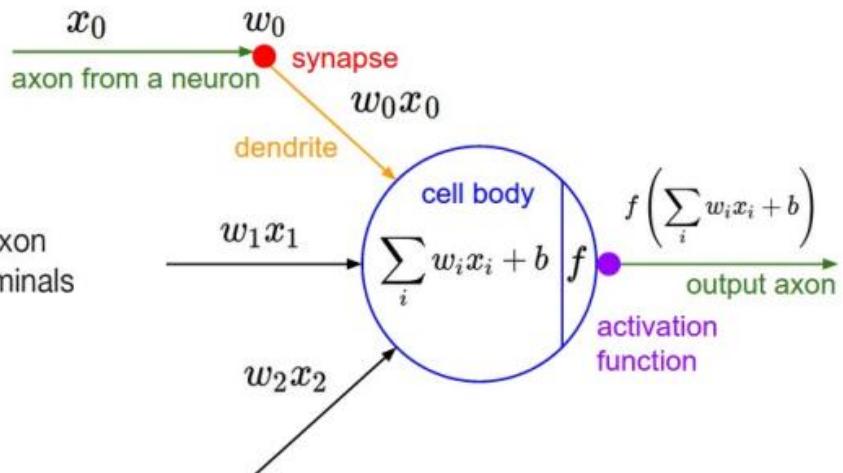
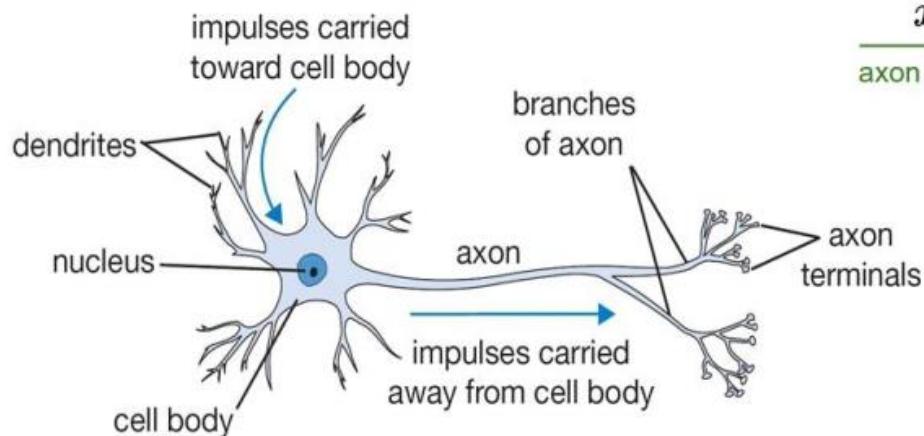
Baidu 百度

Let's get serious



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. The more layers the network has, the higher-level features it will learn.

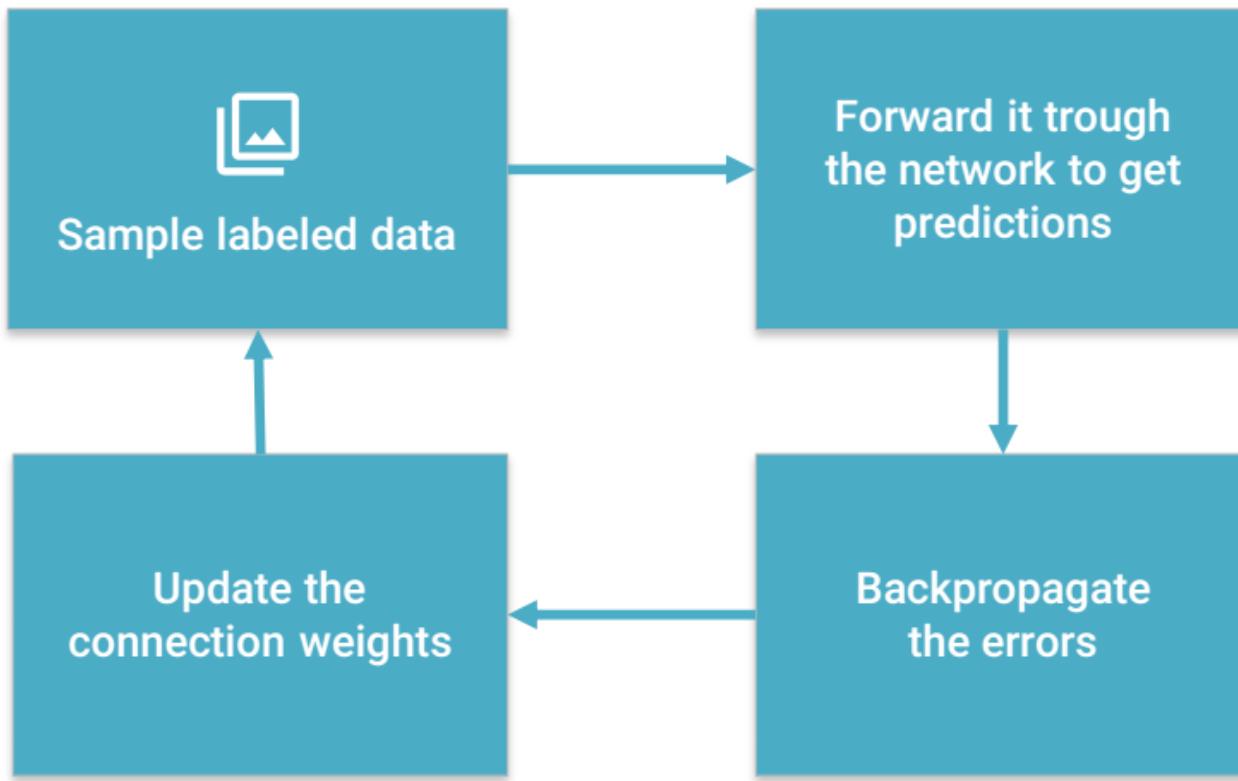
Artificial neurons are somewhat similar to real neurons



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

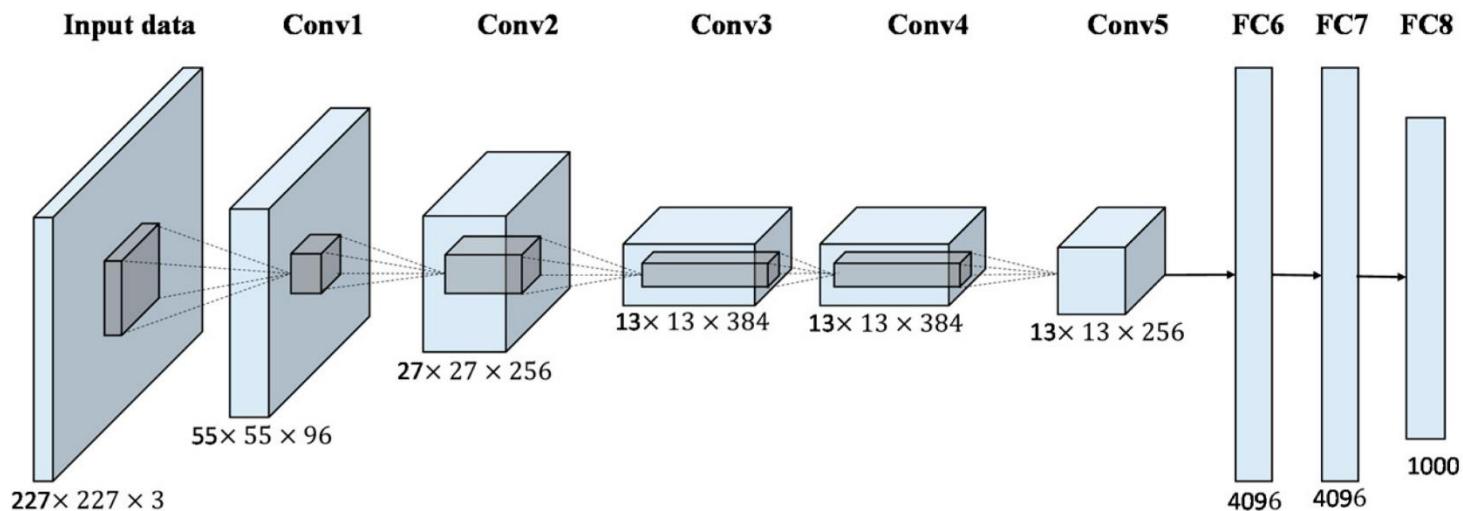


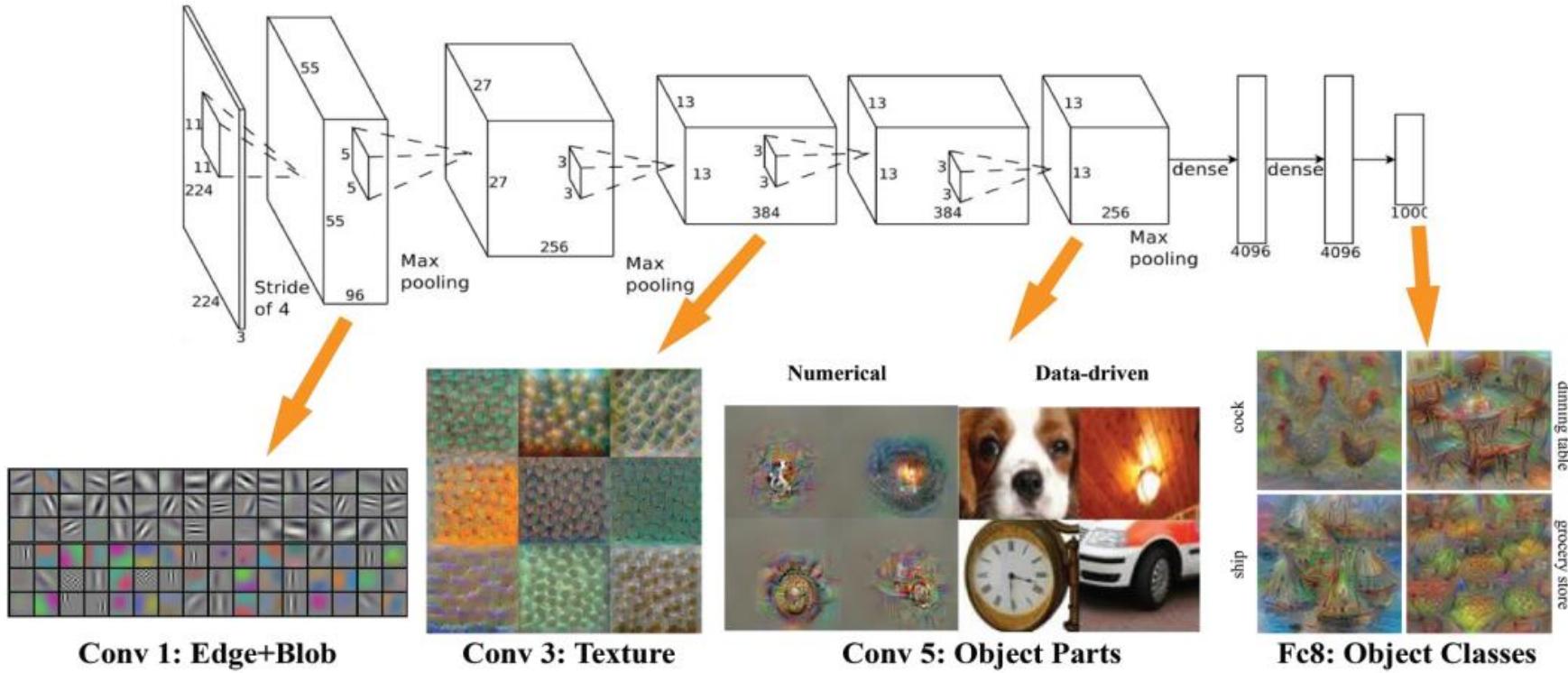
Neurons are **trained to filter and detect specific features** or patterns (e.g. edge, nose) by receiving weighted input, transforming it with the activation function and passing it to the outgoing connections.



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.

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.



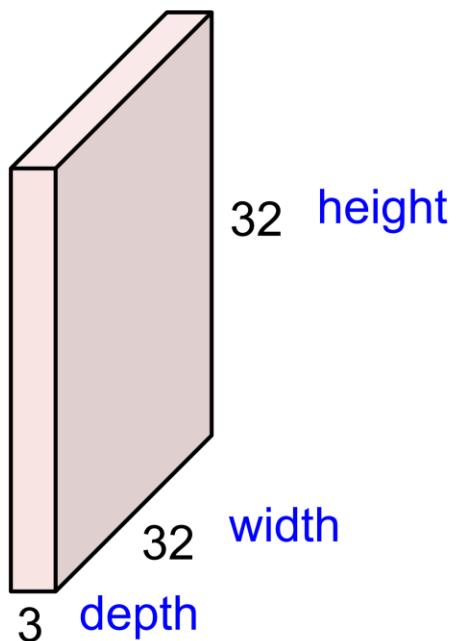


Convolution layer is a feature detector that automagically learns to filter out not needed information from an input by using convolution kernel.

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.

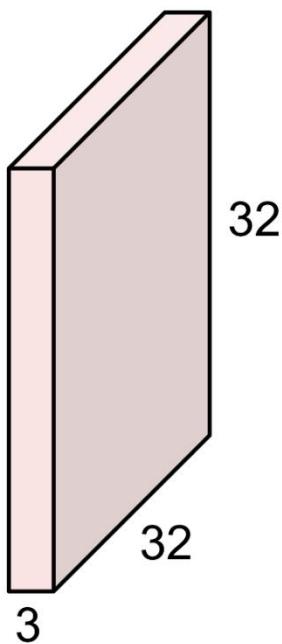
Convolution Layer

32x32x3 image

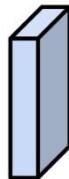


Convolution Layer

32x32x3 image

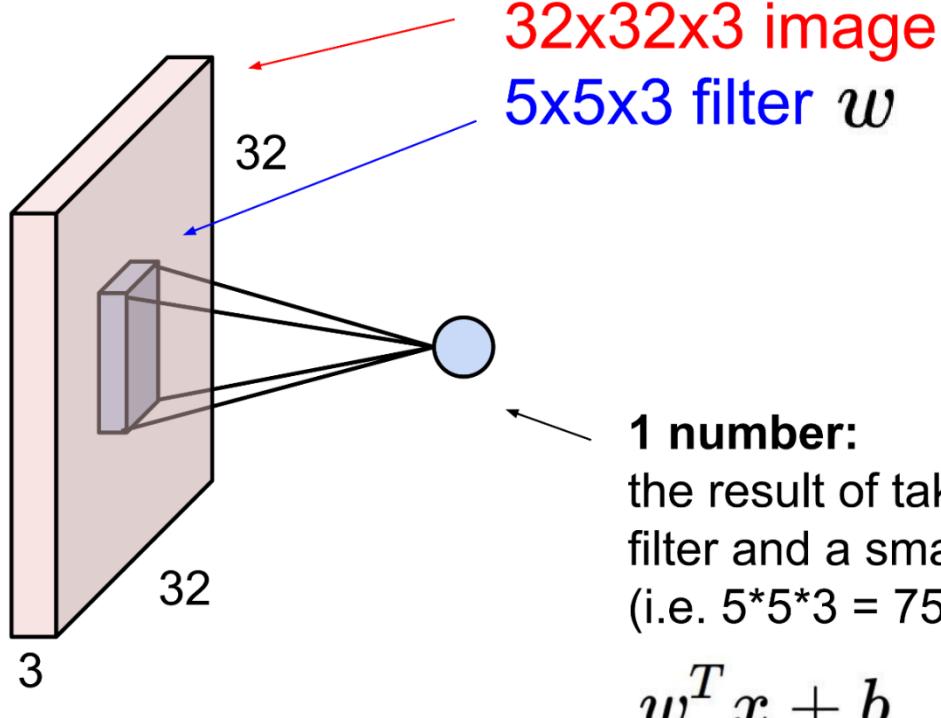


5x5x3 filter

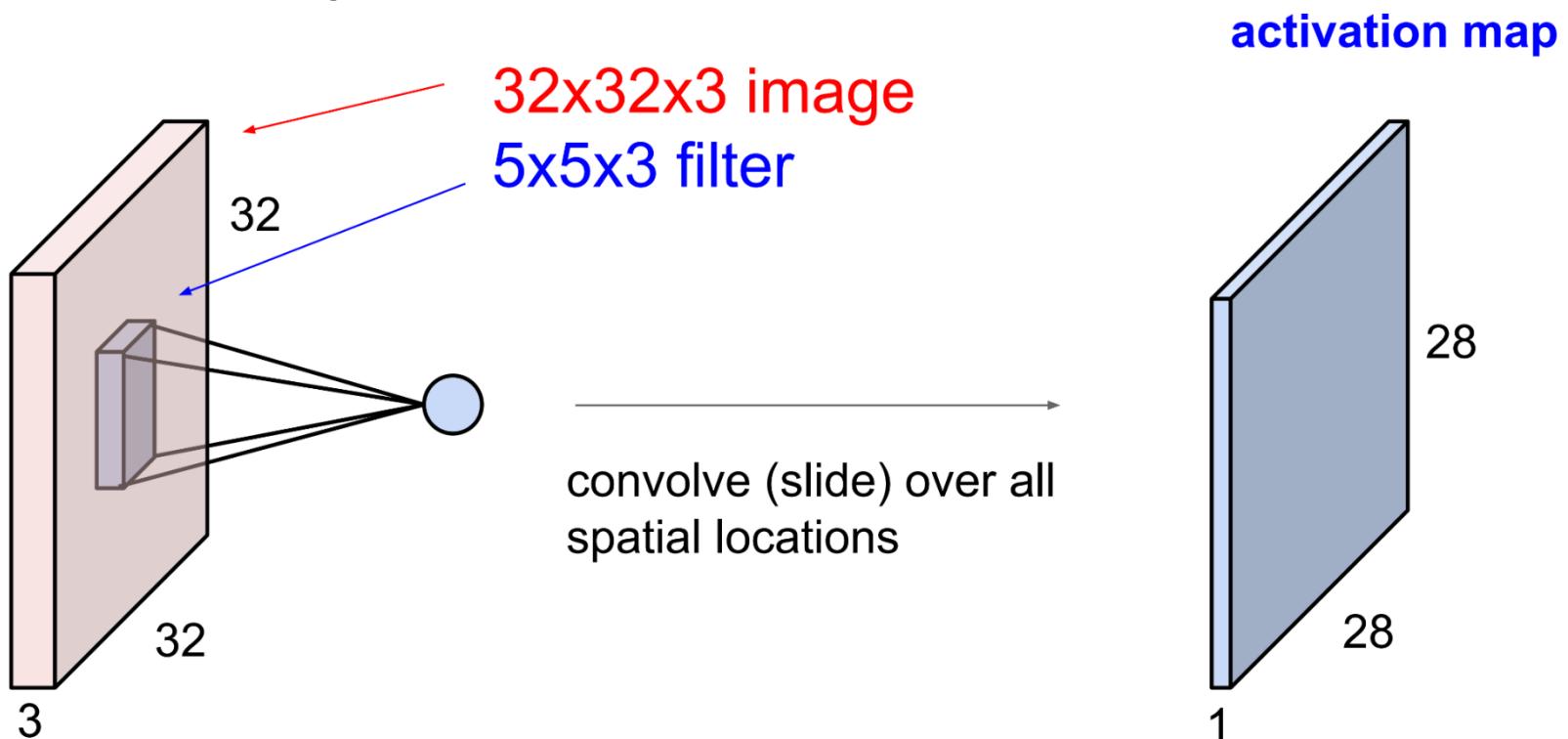


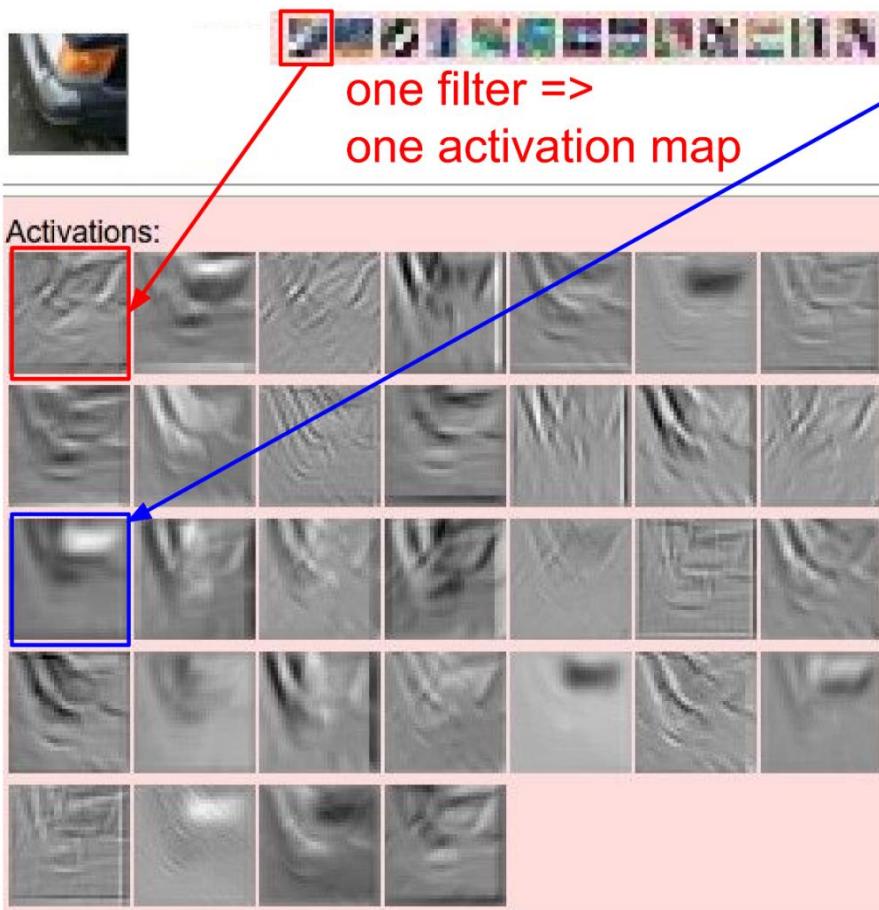
Convolve the filter with the image
i.e. “slide over the image spatially,
computing dot products”

Convolution Layer



Convolution Layer





**example 5x5 filters
(32 total)**

We call the layer convolutional because it is related to convolution of two signals:

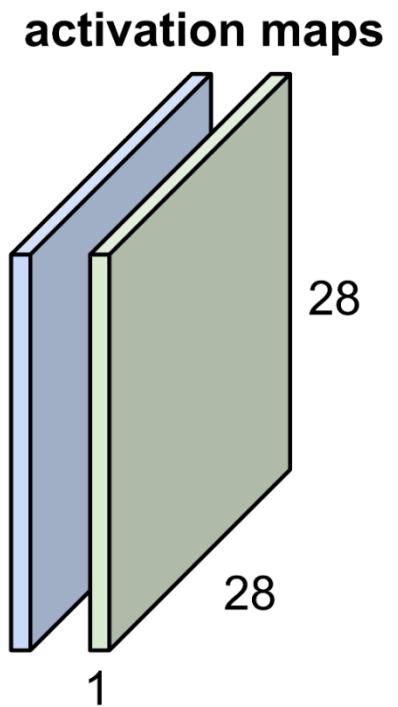
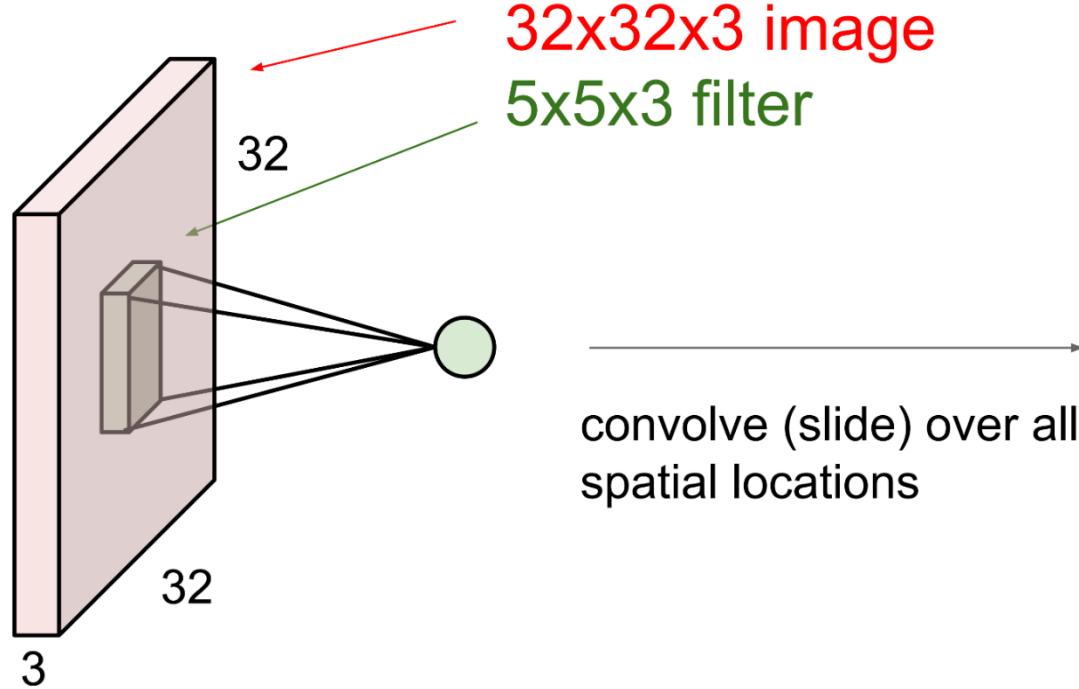
$$f[x,y] * g[x,y] = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f[n_1, n_2] \cdot g[x - n_1, y - n_2]$$



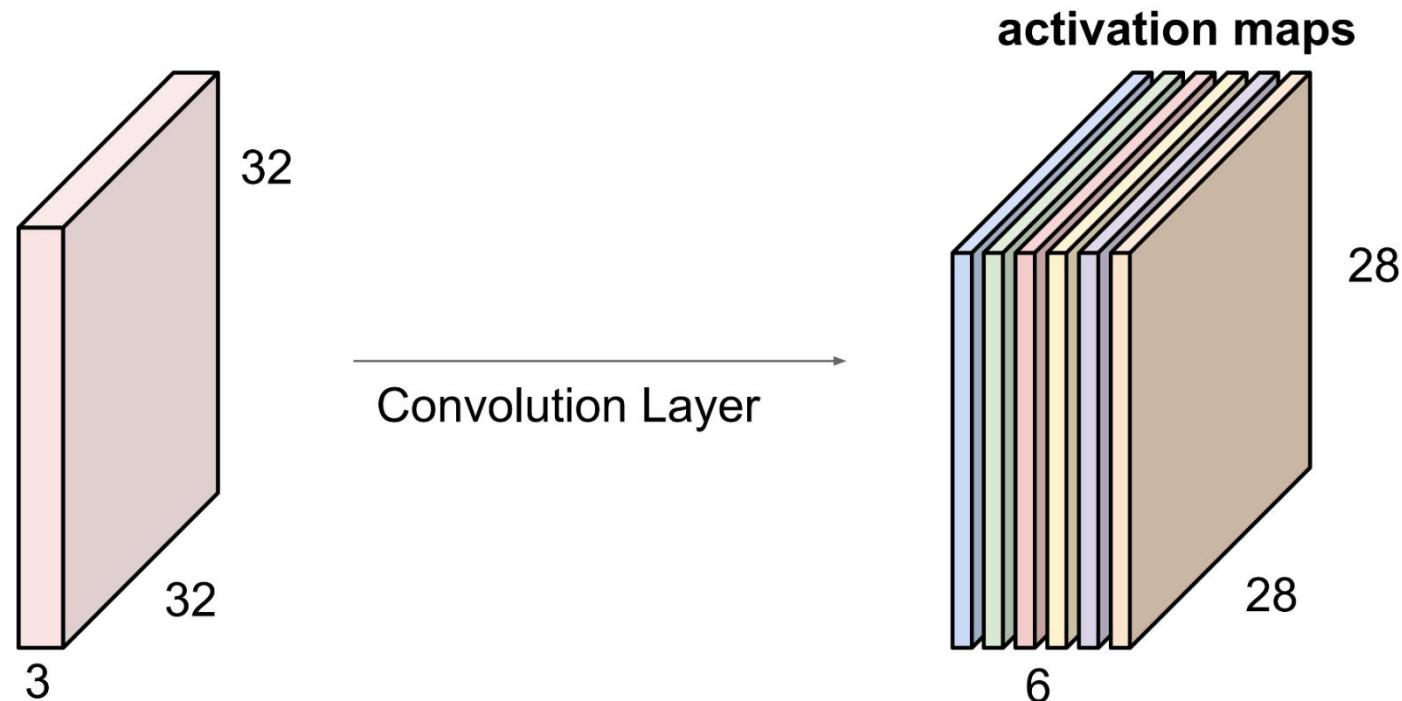
elementwise multiplication and sum of a filter and the signal (image)

Convolution Layer

consider a second, green filter

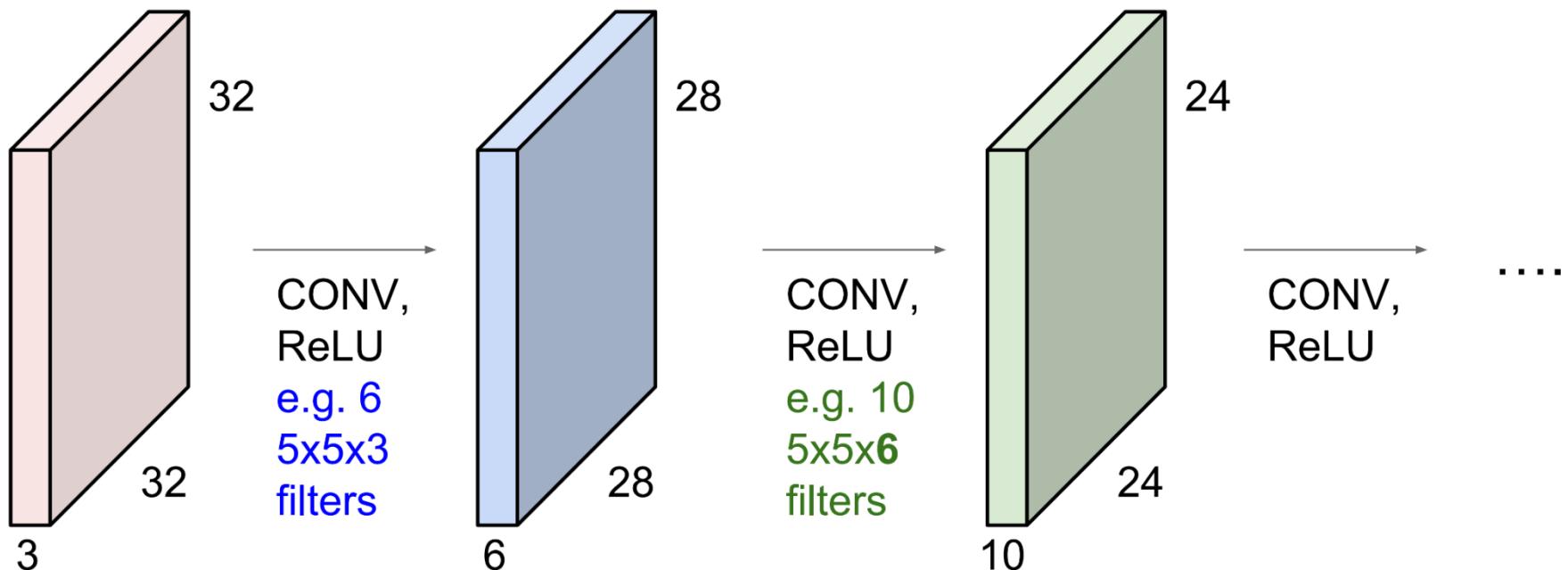


For example, if we had 6 5x5 filters, we'll get 6 separate activation maps:



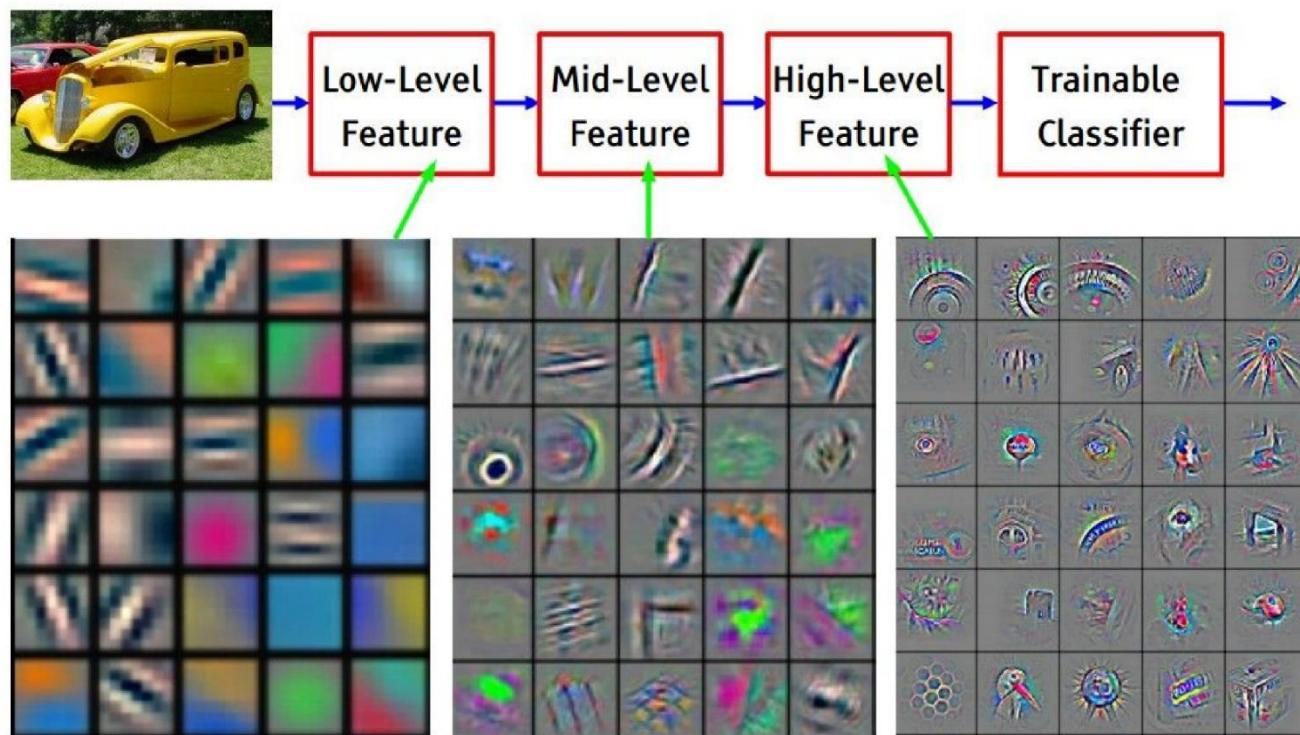
We stack these up to get a “new image” of size 28x28x6!

Preview: ConvNet is a sequence of Convolutional Layers, interspersed with activation functions

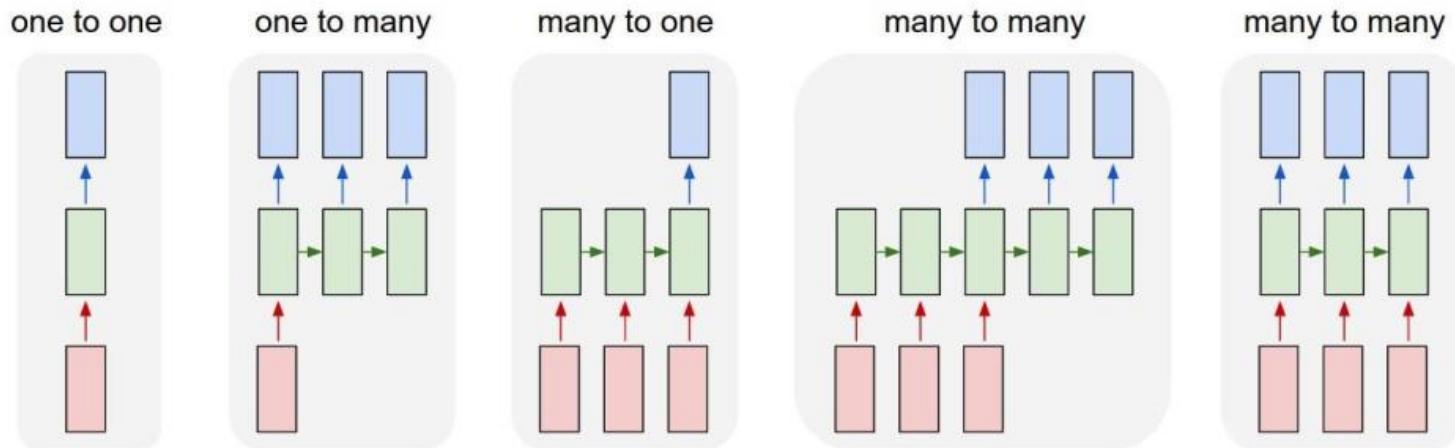


Preview

[From recent Yann LeCun slides]



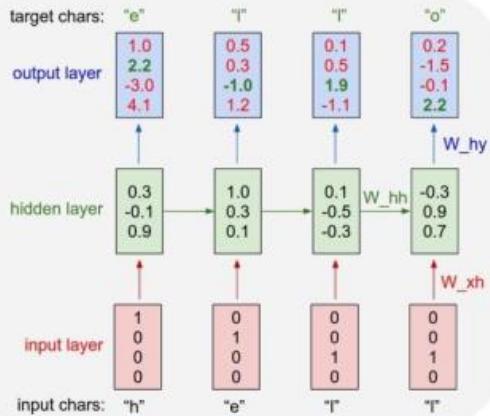
Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]



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.

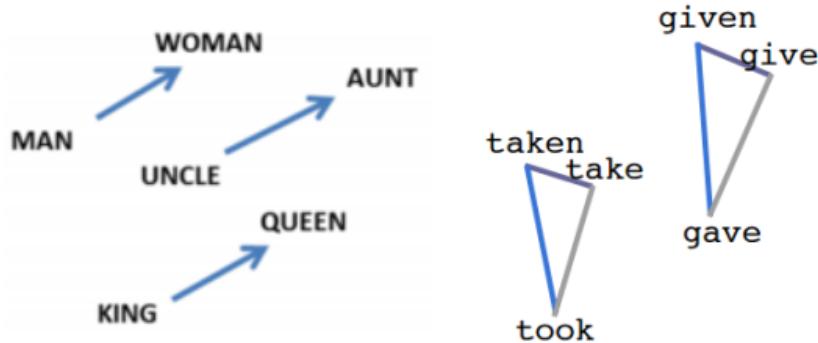


To train the RNN, insert characters sequentially and predict the probabilities of the next letter. Backpropagate error and update RNN's weights to increase the confidence of the correct letter (green) and decrease the confidence of all other letters (red).

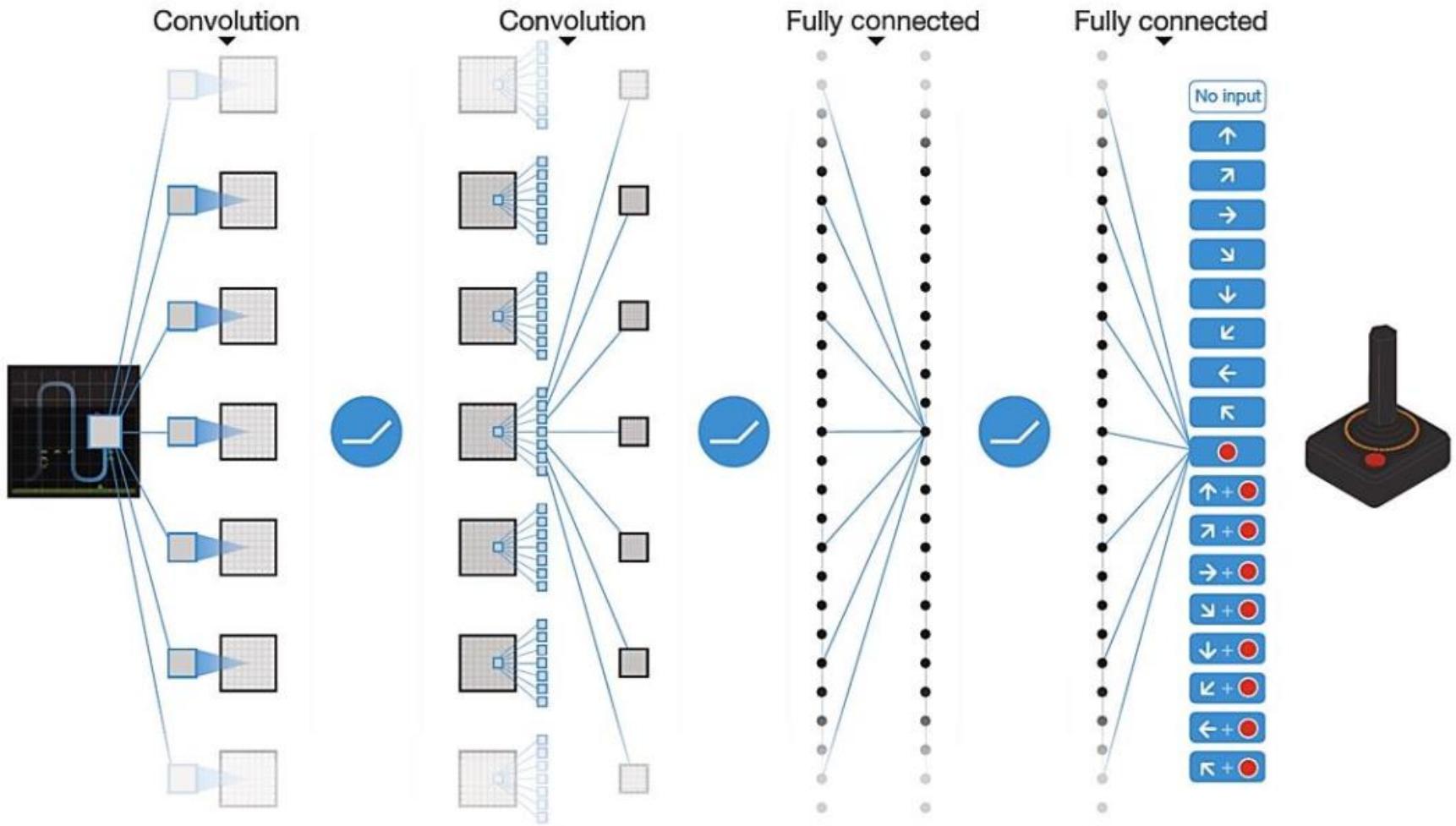
The emperor travelled back to [[Antioch, Perth, October 25|21]] to note, the Kingdom of Costa Rica, unsuccessful fashioned the [[Thrales]], [[Cynth's Dajoard]], known in western [[Scotland]], near Italy to the conquest of India with the conflict. Copyright was the succession of independence in the slop of Syrian influence that was a famous German movement based on a more popular servitious, non-doctrinal and sexual power post. Many governments recognize the military housing of the [[Civil Liberalization and Infantry Resolution 265 National Party in Hungary]], that is sympathetic to be to the [[Punjab Resolution]] (PJS)[<http://www.humah.yahoo.com/guardian.cfm/7754800786d17551963s89.htm>]

Trained on structured Wikipedia markdown. Network learns to spell English words completely from scratch and copy general syntactic structures.

| 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 |

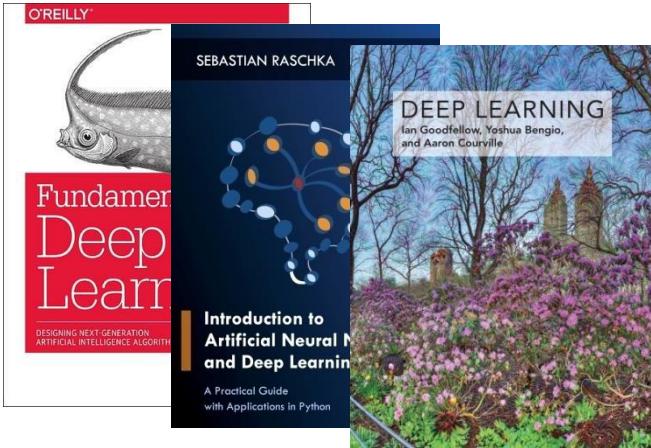


Woman – Man ≈ Aunt - Uncle
 King - Male + Female ≈ Queen
 Human - Animal ≈ Ethics



Deep Q-Learning (DQN) is a model-free approach to reinforcement learning using deep networks in environments with discrete action choices

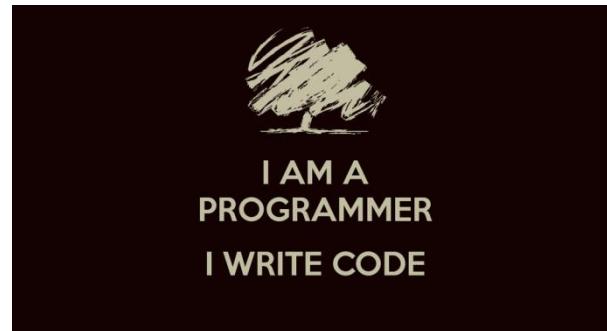
Ok, I'm sold, what should I do next (besides taking Machine Learning courses)



**Read about it
(serious reading)**



MOOCs



Most frameworks are open source

Caffe



DL4J
Deeplearning4j

MINERVA

mxnet



MatConvNet



theano



Great team, great professors, great research



Domingo Mery



Álvaro Soto



Jorge Baier



Karim Pichara



Cristian Ruz



Denis Parra



Hans Löbel

Great team, great professors, great research



IA Lab



Takeaways

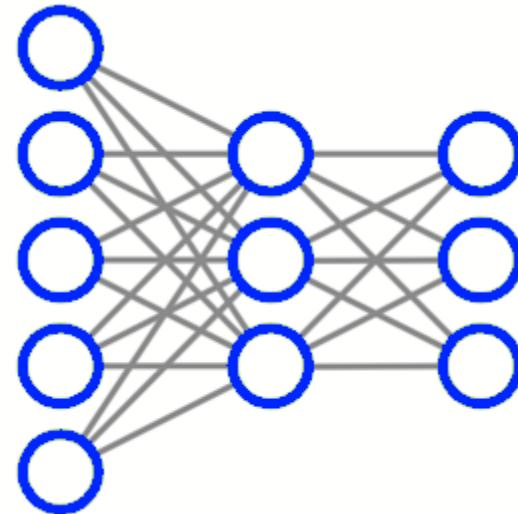
- Deep Learning is not magic! Just statistics and optimization in a black box.
- Exceptionally effective at recognizing patterns.
- Still no common sense (can be even racist).
- If you're interested, learn it now.
- If not, learn it anyways, at least the basics.
- This is one of the technologies that will change the world.

Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ciencia de la Computación



IIC1005 – Computación: Ciencia y Tecnología del Mundo Digital

Introducción a Deep Learning



Hans Löbel