

# Introduction to deep learning

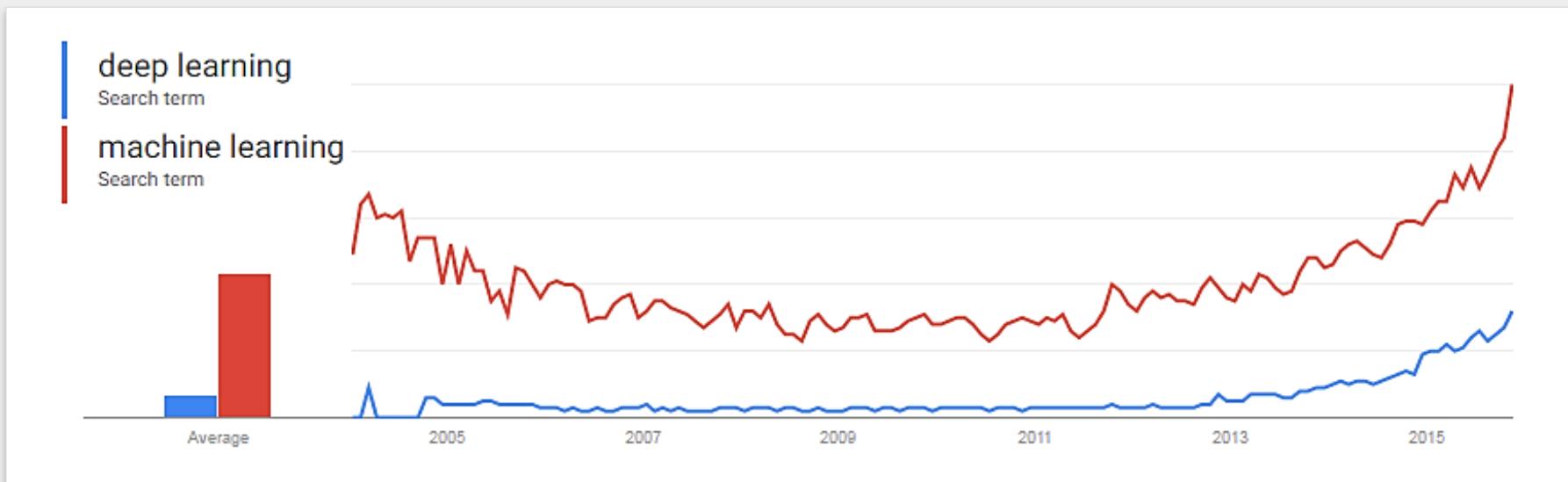
Cours M2 AIDN

**Nicolas Courty**

Professeur UBS, Laboratoire IRISA  
[courty@univ-ubs.fr](mailto:courty@univ-ubs.fr)

# Interest

Google NGRAM & Google Trends



# Hype or Reality?

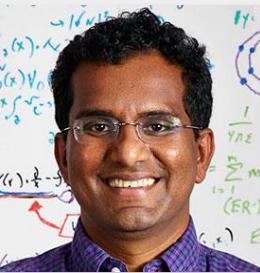
## Quotes



I have worked all my life in Machine Learning, and I've never seen one algorithm knock over benchmarks like Deep Learning  
– Andrew Ng (*Stanford & Baidu*)



Deep Learning is an algorithm which has no theoretical limitations of what it can learn; the more data you give and the more computational time you provide, the better it is – Geoffrey Hinton (*Google*)



Human-level artificial intelligence has the potential to help humanity thrive more than any invention that has come before it – Dileep George (*Co-Founder Vicarious*)



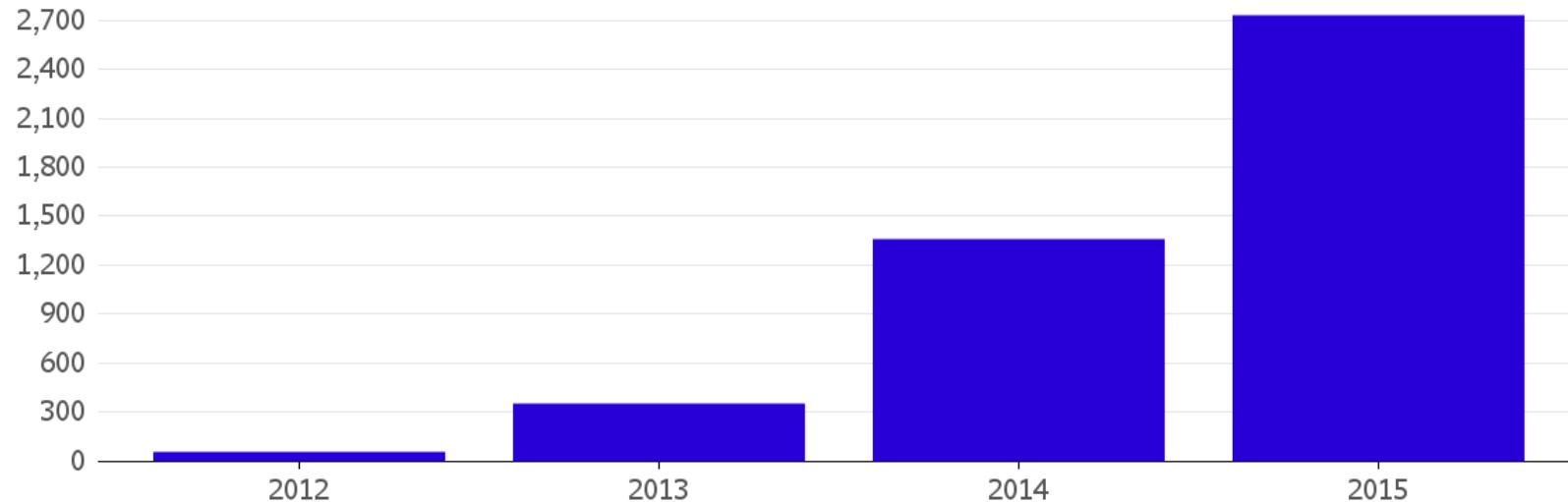
For a very long time it will be a complementary tool that human scientists and human experts can use to help them with the things that humans are not naturally good – Demis Hassabis (*Co-Founder DeepMind*)

# Hype or Reality?

## Deep Learning at Google

### Artificial Intelligence Takes Off at Google

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



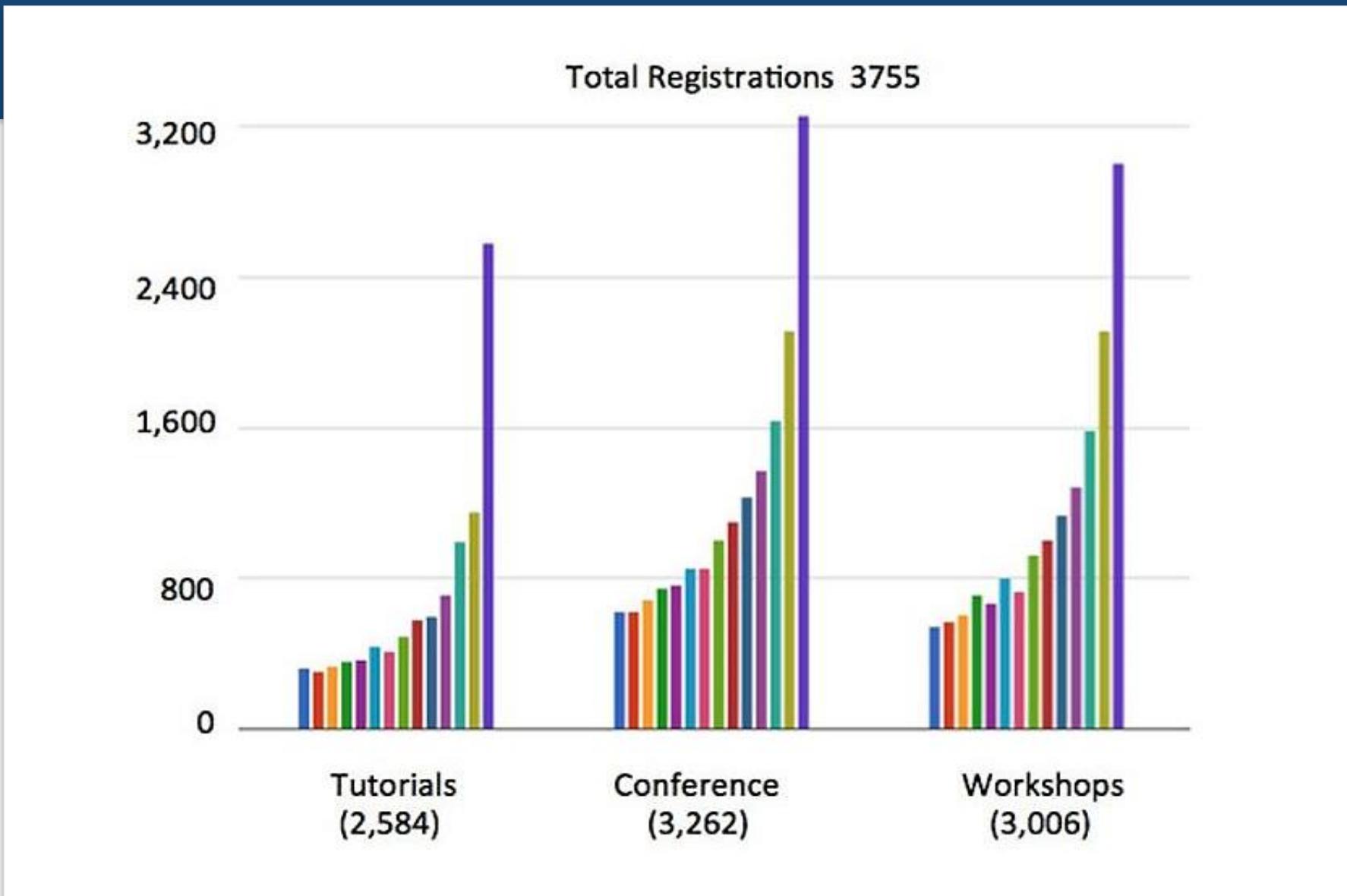
Source: Google

Note: 2015 data does not incorporate data from Q4

Bloomberg

# Hype or Reality?

NIPS (Computational Neuroscience Conference) Growth



# What is Artificial Intelligence?

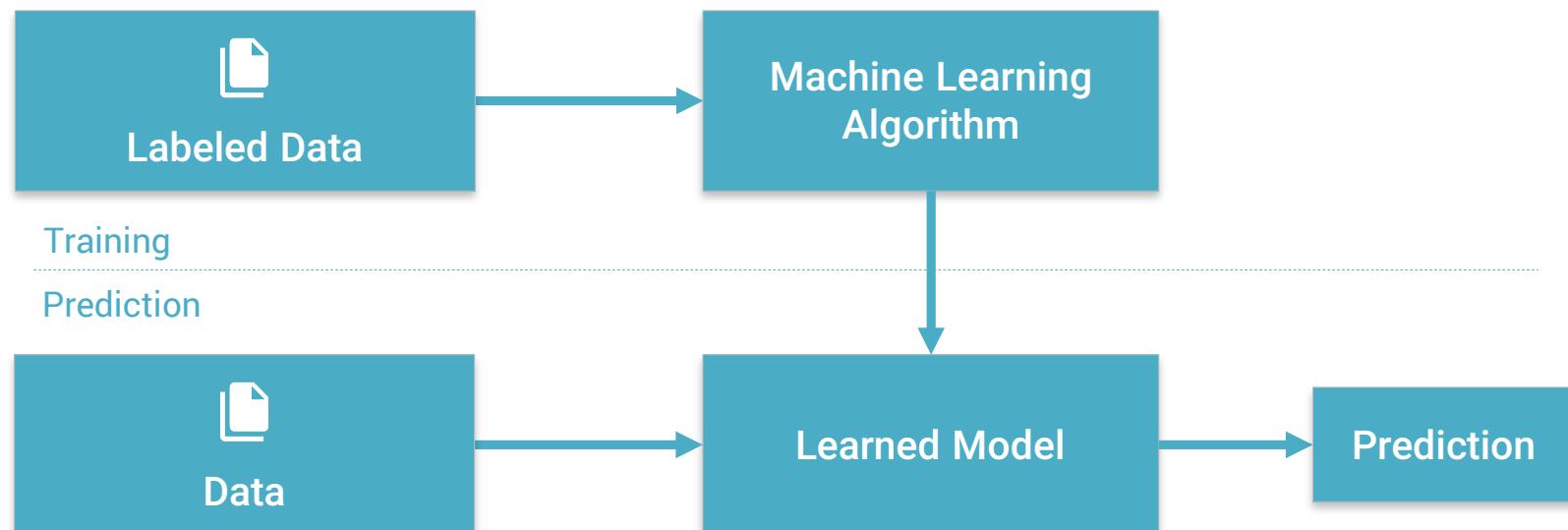


# Machine Learning - Basics

## Introduction



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

# Machine Learning - Basics

## Learning Approaches



Supervised Learning: Learning with a **labeled training set**

*Example: email spam detector with training set of already labeled emails*



Unsupervised Learning: **Discovering patterns** in unlabeled data

*Example: cluster similar documents based on the text content*

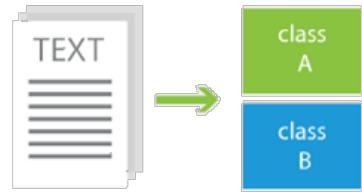


Reinforcement Learning: learning based on **feedback** or reward

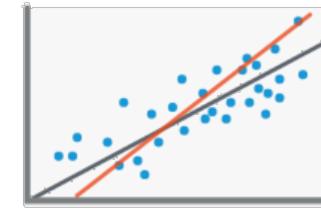
*Example: learn to play chess by winning or losing*

# Machine Learning - Basics

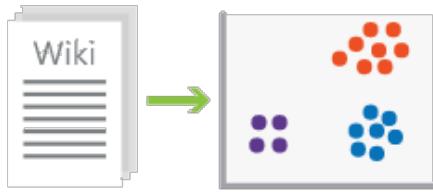
## Problem Types



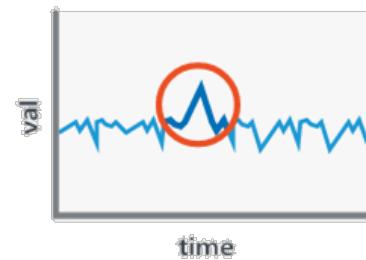
Classification  
(supervised – predictive)



Regression  
(supervised – predictive)



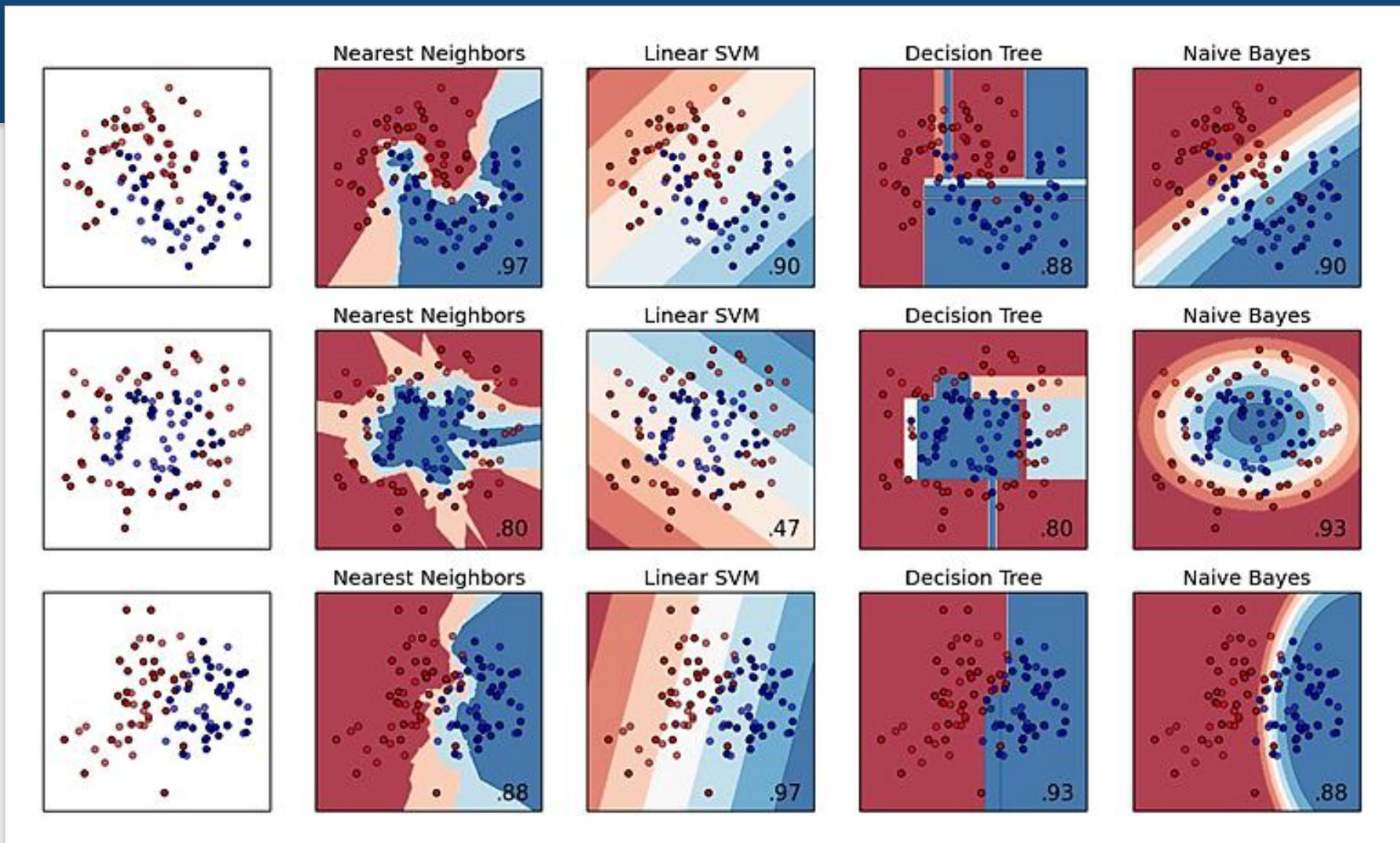
Clustering  
(unsupervised – descriptive)



Anomaly Detection  
(unsupervised – descriptive)

# Machine Learning - Basics

## Algorithms Comparison - Classification



Scikit-Learn

# What is Deep Learning?



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

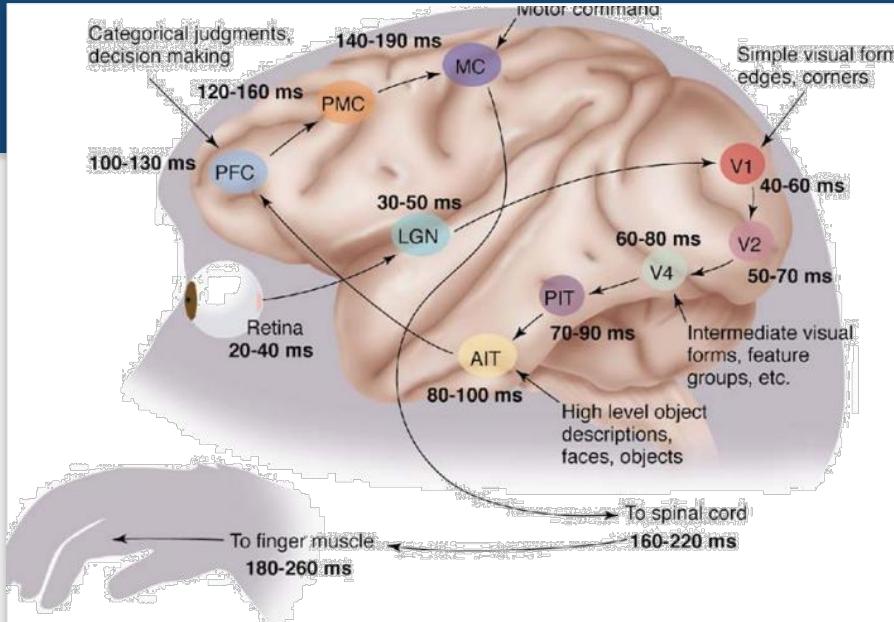


Utilizes learning algorithms that derive meaning out of data by using a **hierarchy** of multiple layers that **mimic the neural networks of our brain**.



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

# Inspired by the Brain



The first **hierarchy of neurons** that receives information in the visual cortex are sensitive to specific edges while brain regions further down the visual pipeline are sensitive to more complex structures such as faces.



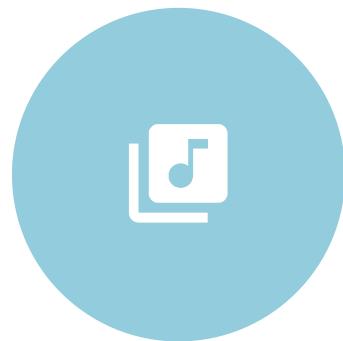
Our brain has lots of neurons connected together and the **strength of the connections** between neurons represents **long term knowledge**.

1

**One learning algorithm hypothesis:** all significant mental algorithms are learned except for the learning and reward machinery itself.

# Why Deep Learning?

## Applications



Speech  
Recognition



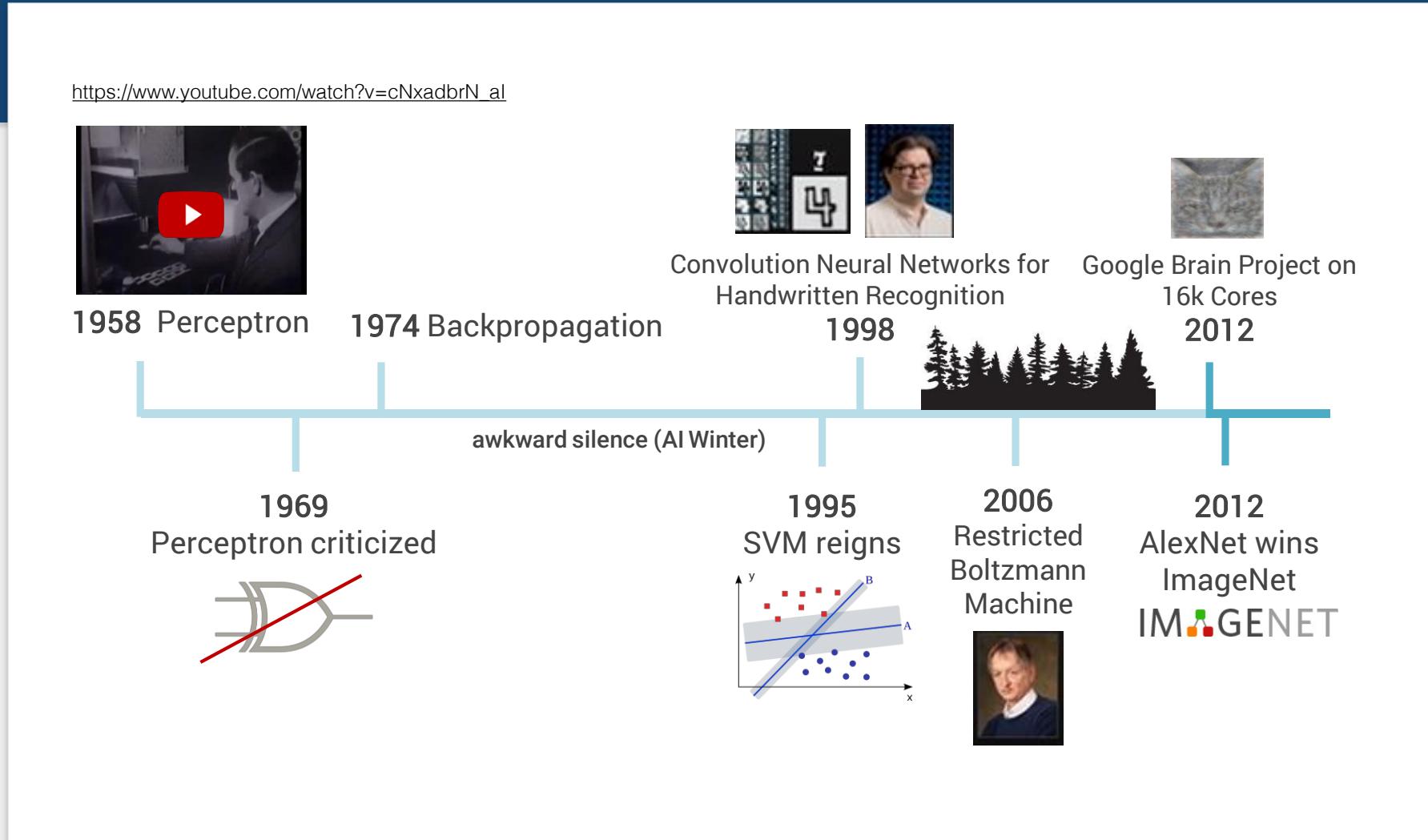
Computer  
Vision



Natural Language  
Processing

# A brief History

A long time ago...



# A brief History

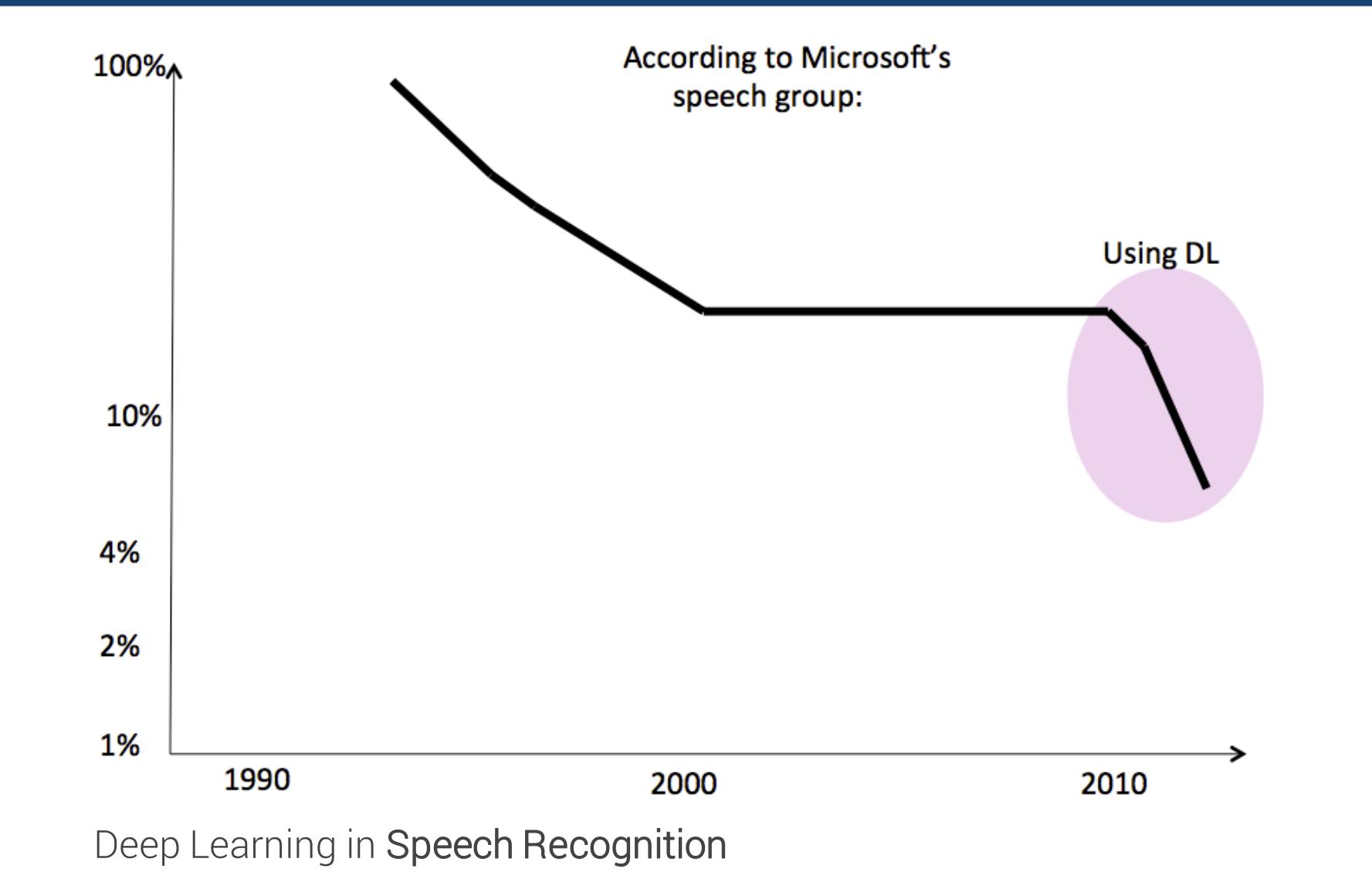
The Big Bang aka “One net to rule them all”



ImageNet: The “computer vision World Cup”

# A brief History

The Big Bang aka “One net to rule them all”

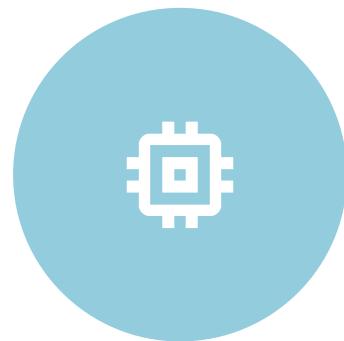


# What changed?

Old wine in new bottles



Big Data  
(Digitalization)



Computation  
(Moore's Law, GPUs)



Algorithmic  
Progress

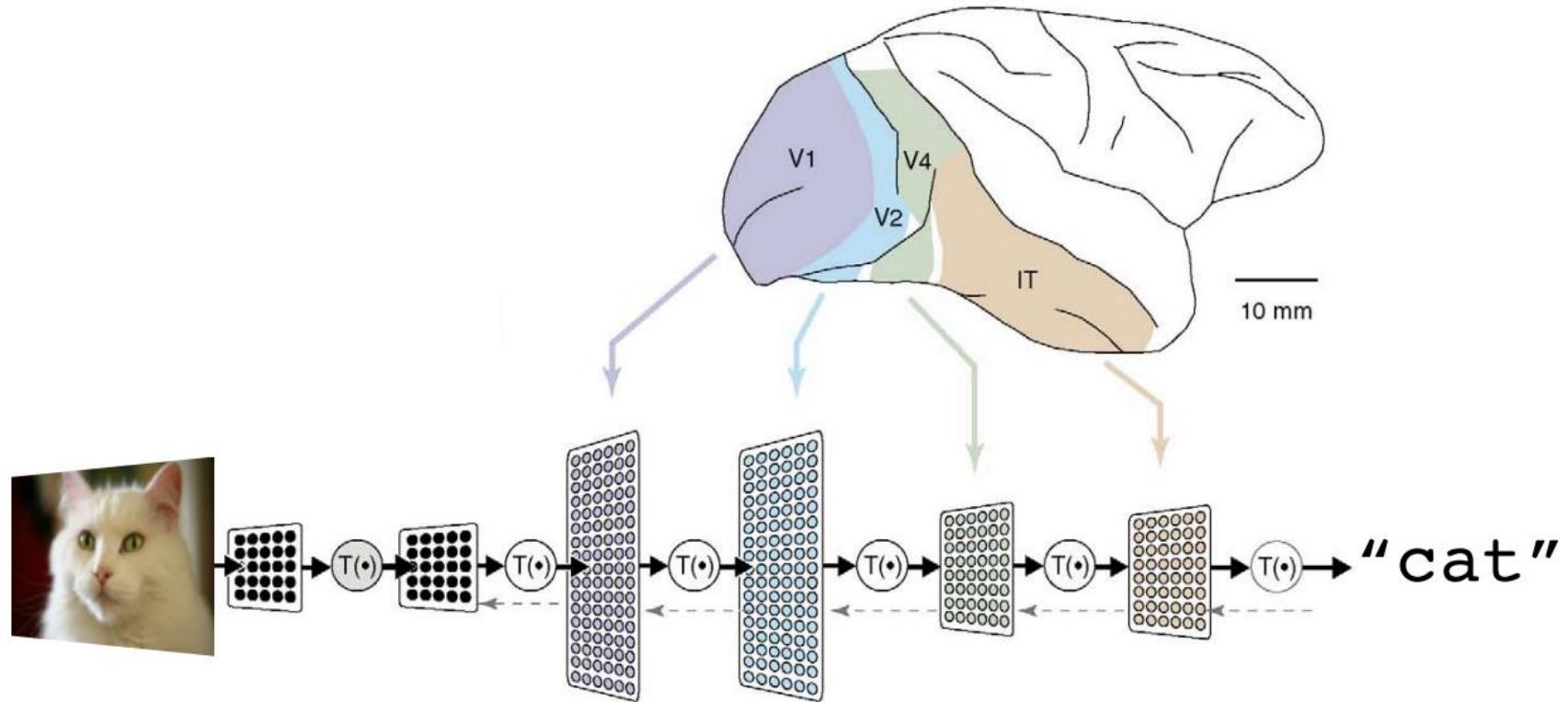
# Deep Learning - Basics

No more feature engineering



# Deep Learning - Basics

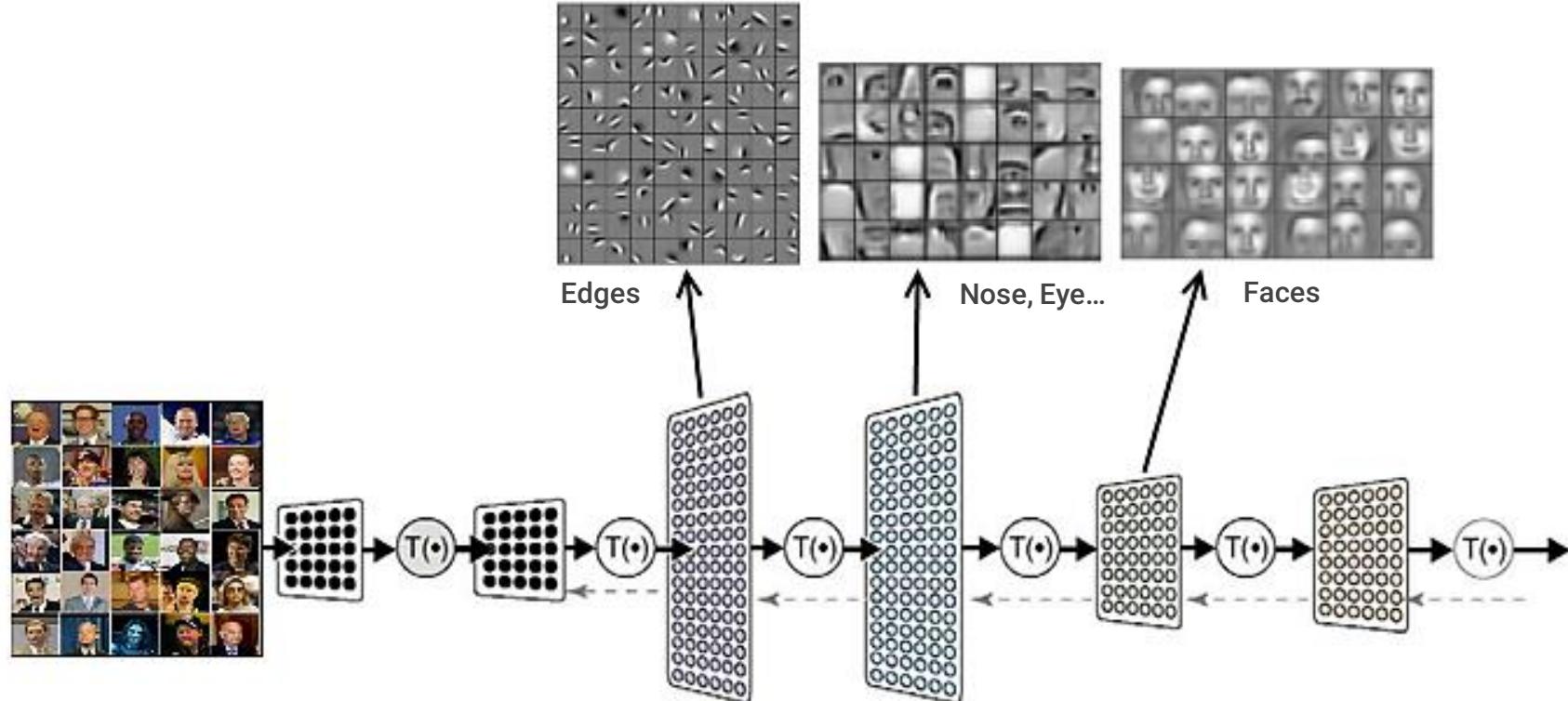
## Architecture



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.

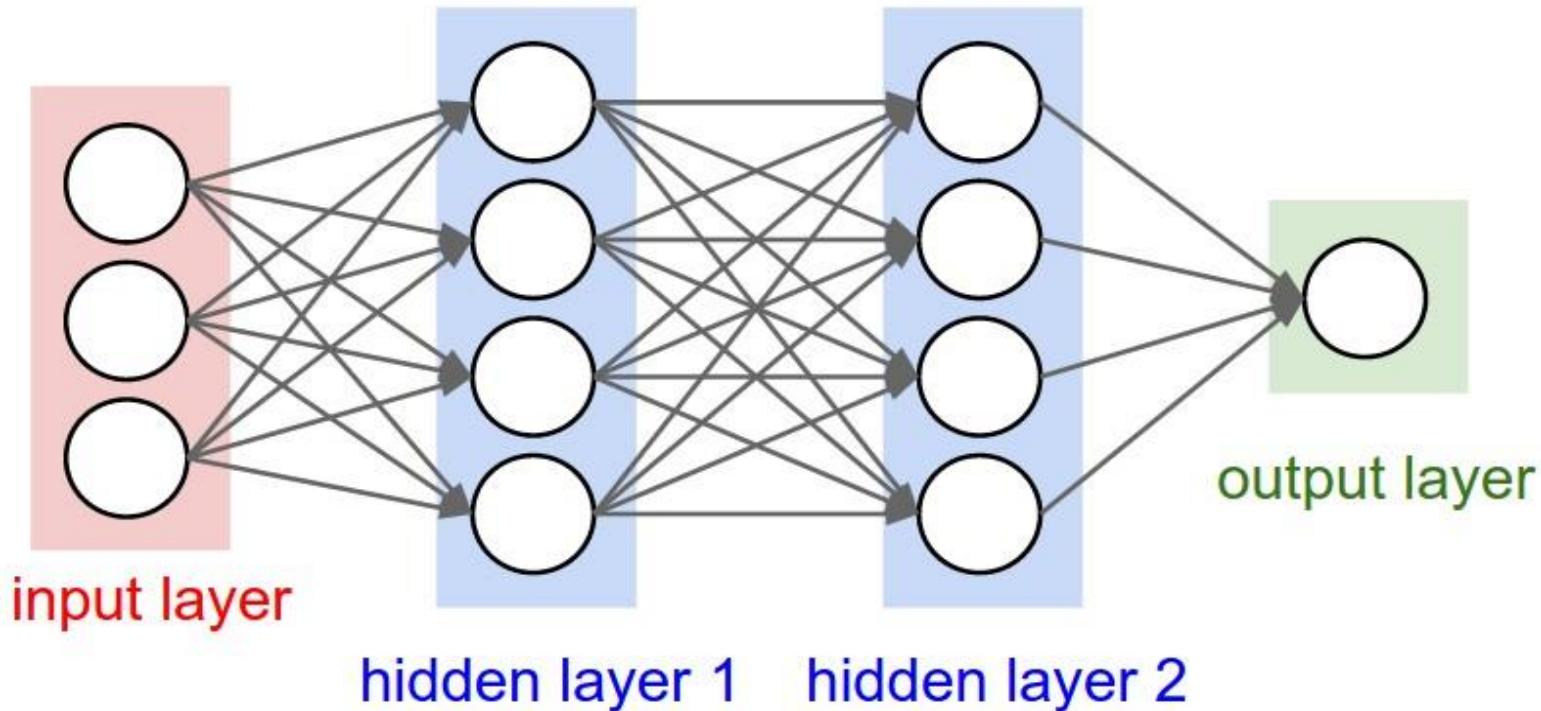
# Deep Learning - Basics

What did it learn?



# Deep Learning - Basics

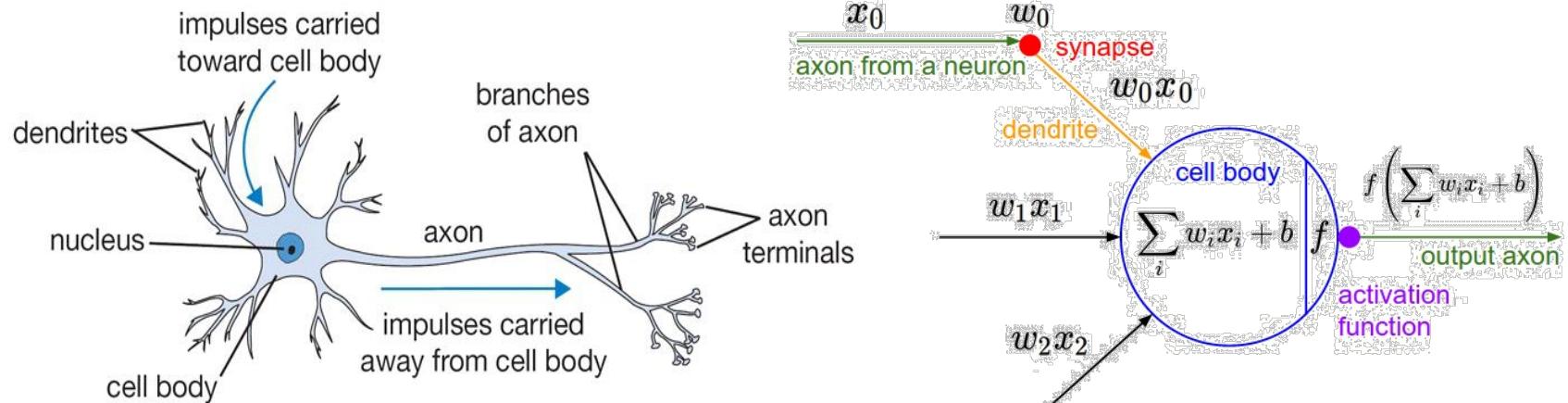
## Artificial Neural Networks



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

# Deep Learning - Basics

## The Neuron



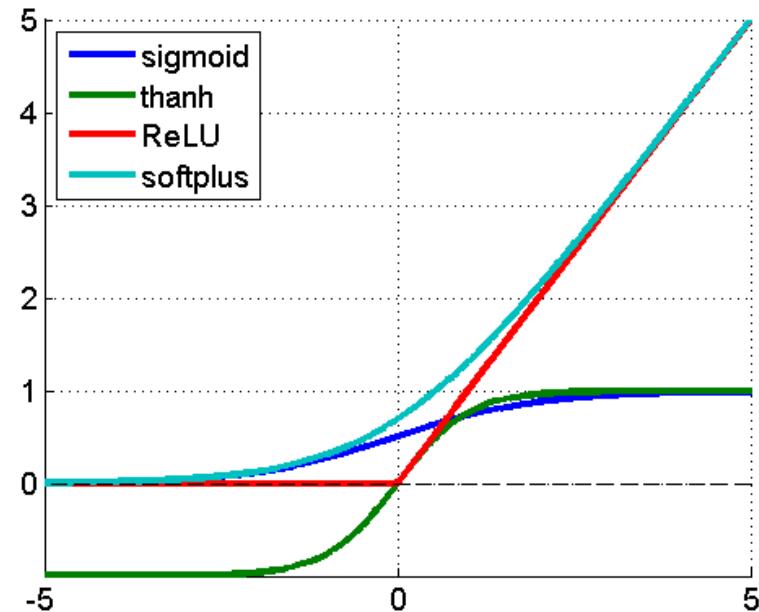
An artificial neuron contains a **nonlinear activation function** and has several incoming and outgoing **weighted connections**.



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 and passing it to the outgoing connections.

# Deep Learning - Basics

## Non-linear Activation Function



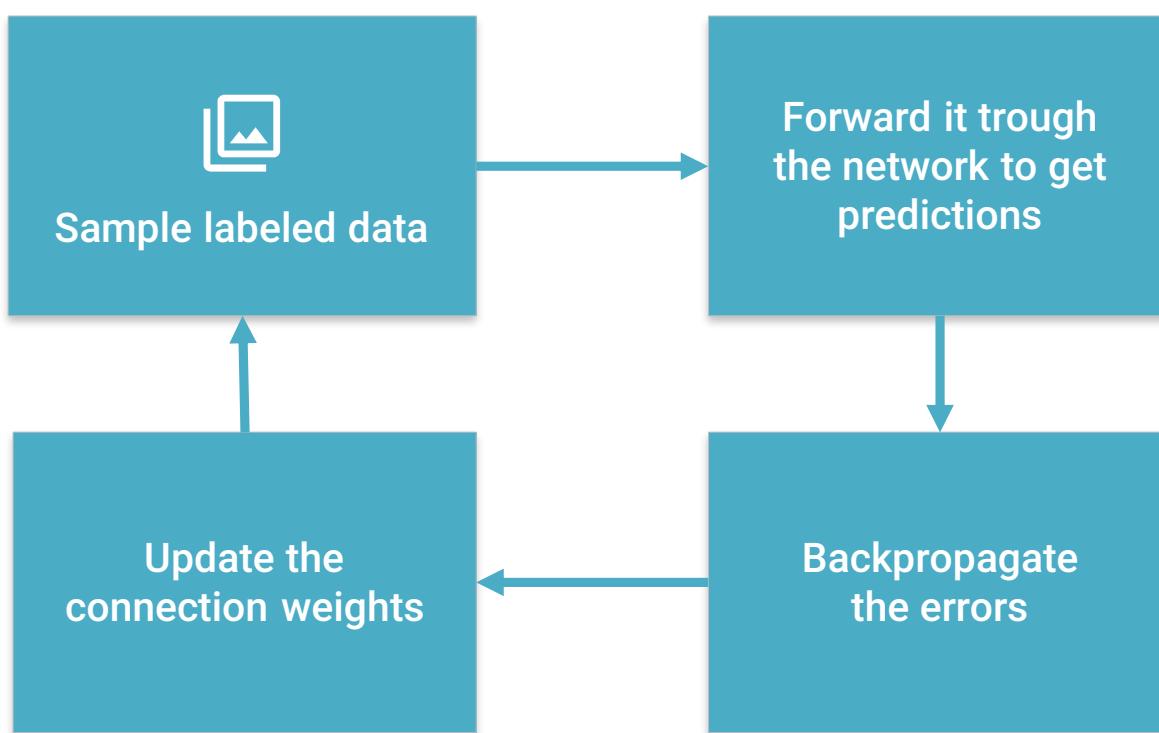
Most deep networks use **ReLU** -  $\max(0, x)$  - nowadays for hidden layers, since it trains much faster, is more expressive than logistic function and prevents the gradient vanishing problem.



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

# Deep Learning - Basics

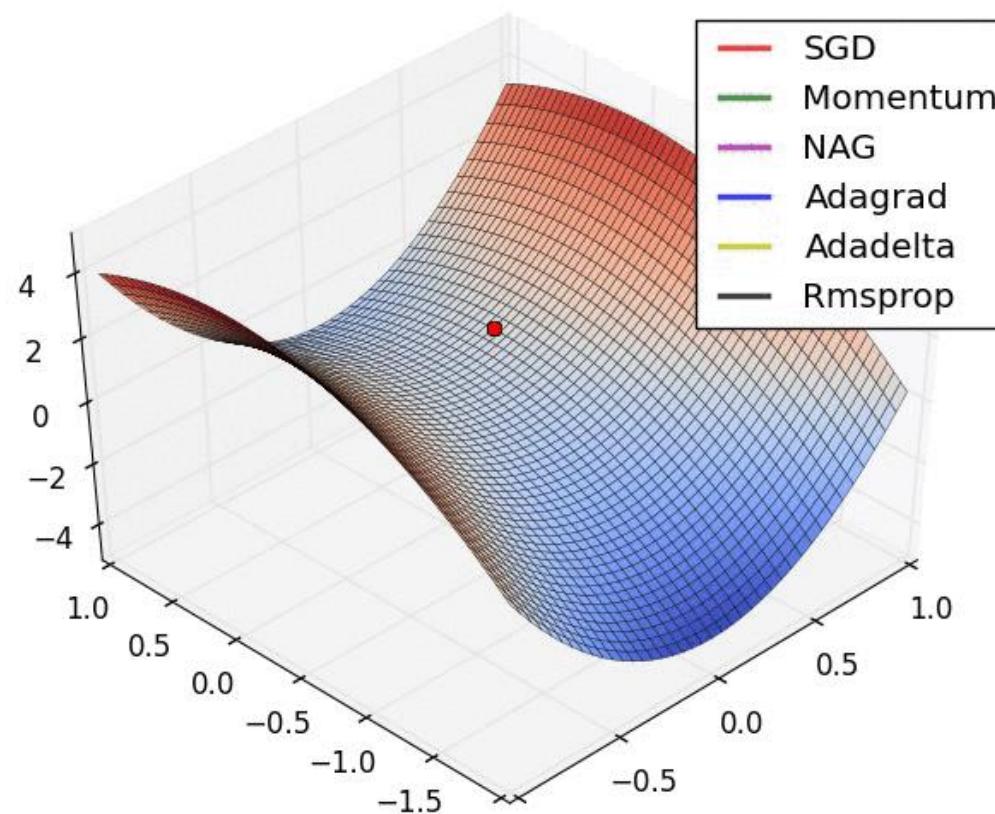
## The Training Process



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.

# Deep Learning - Basics

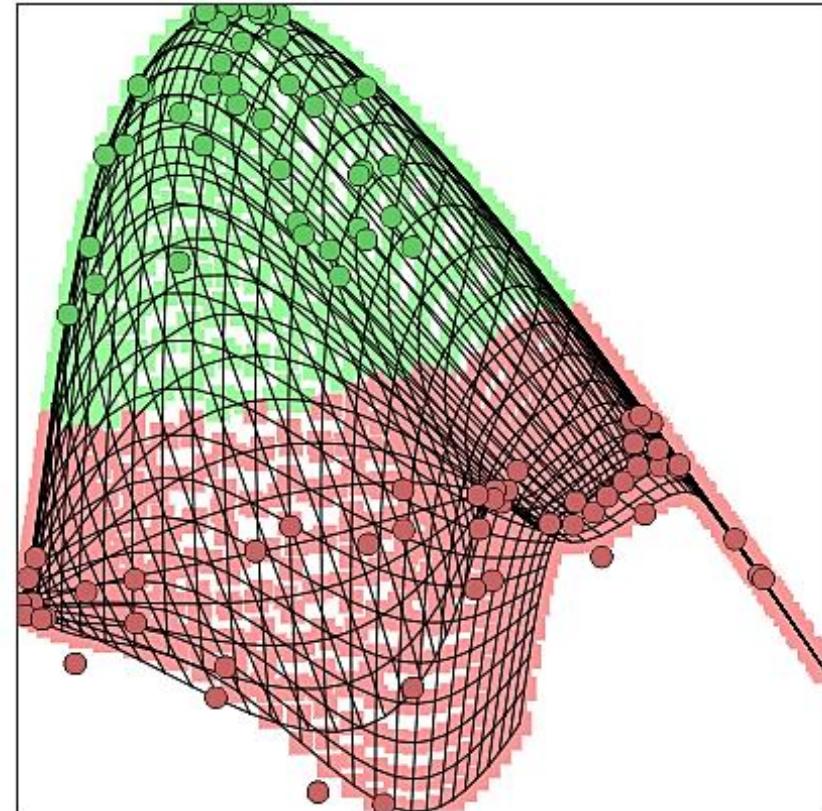
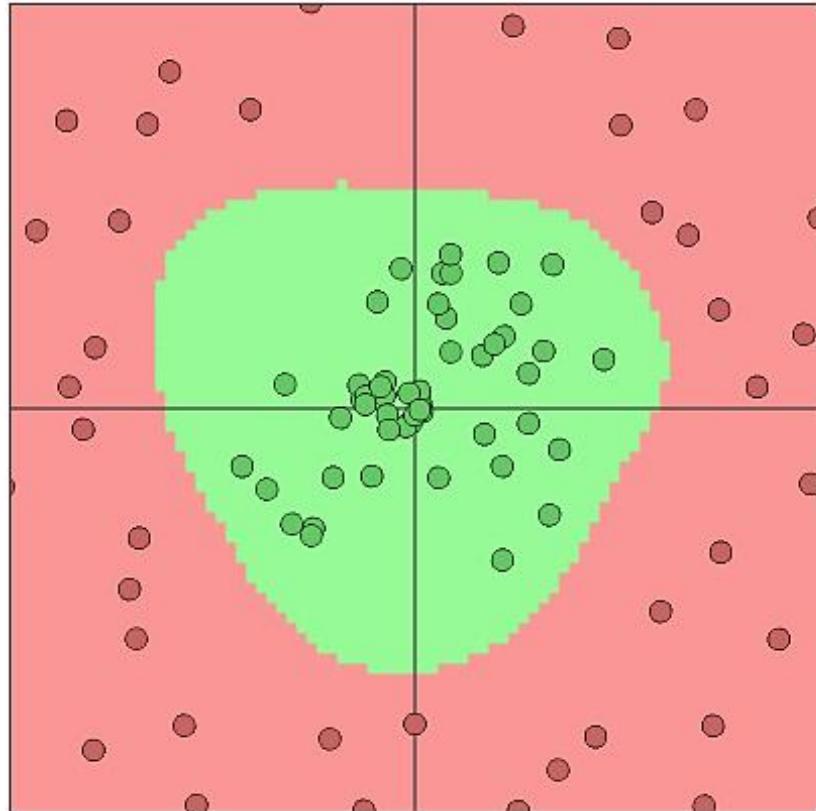
## Gradient Descent



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

# Deep Learning - Basics

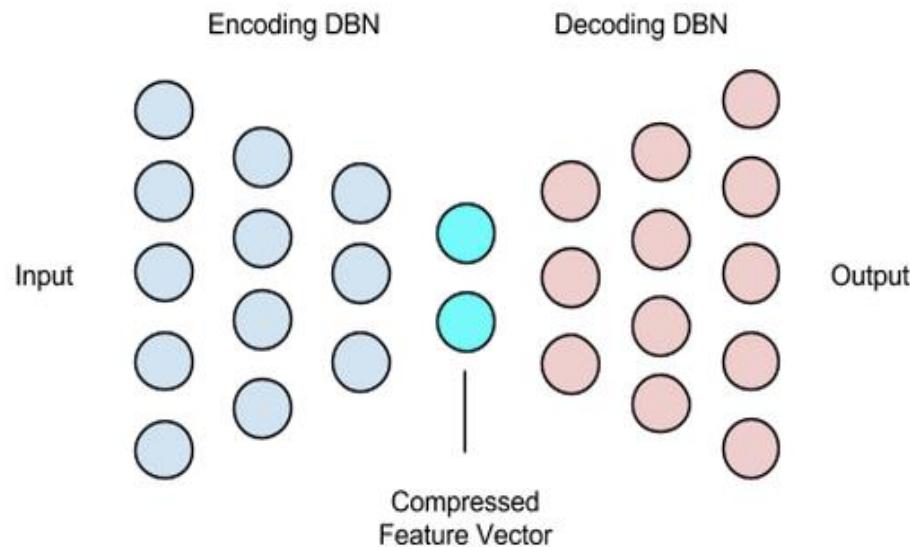
## Data transformation in other dimensions



A neural network is **transforming the data into other dimensions** to solve the specified problem.

# 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.

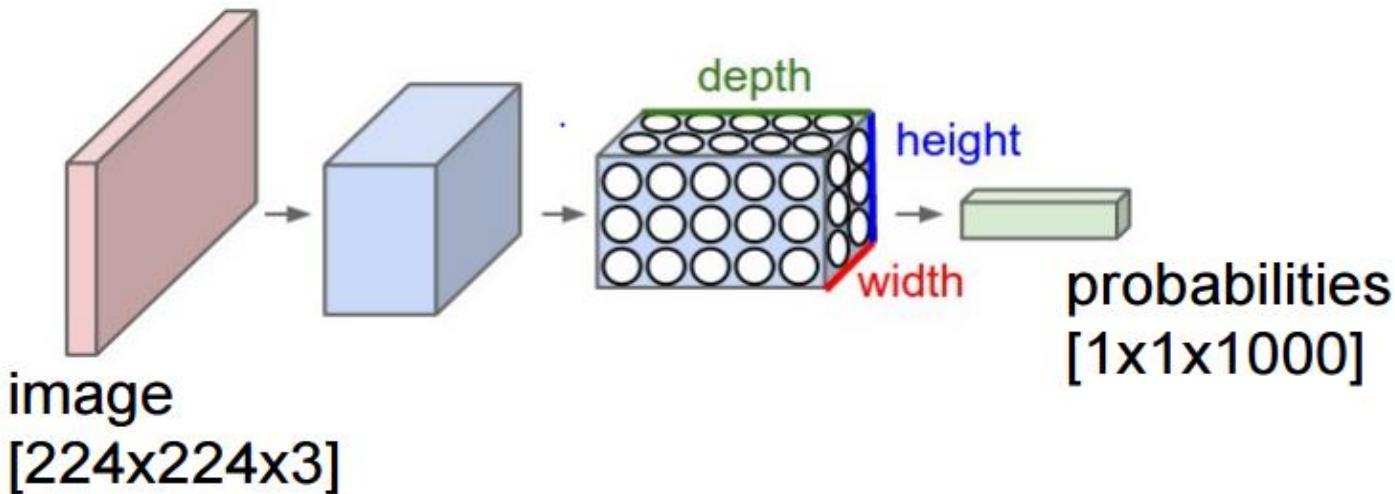


**Topic Modeling:** Document in a collection is converted to a Bag-of-Words and transformed to a compressed feature vector using an autoencoder. The distance from every other document-vector can be measured and nearby document-vectors fall under the same topic.

# Deep Learning - Basics

## Convolutional Neural Nets (CNN)

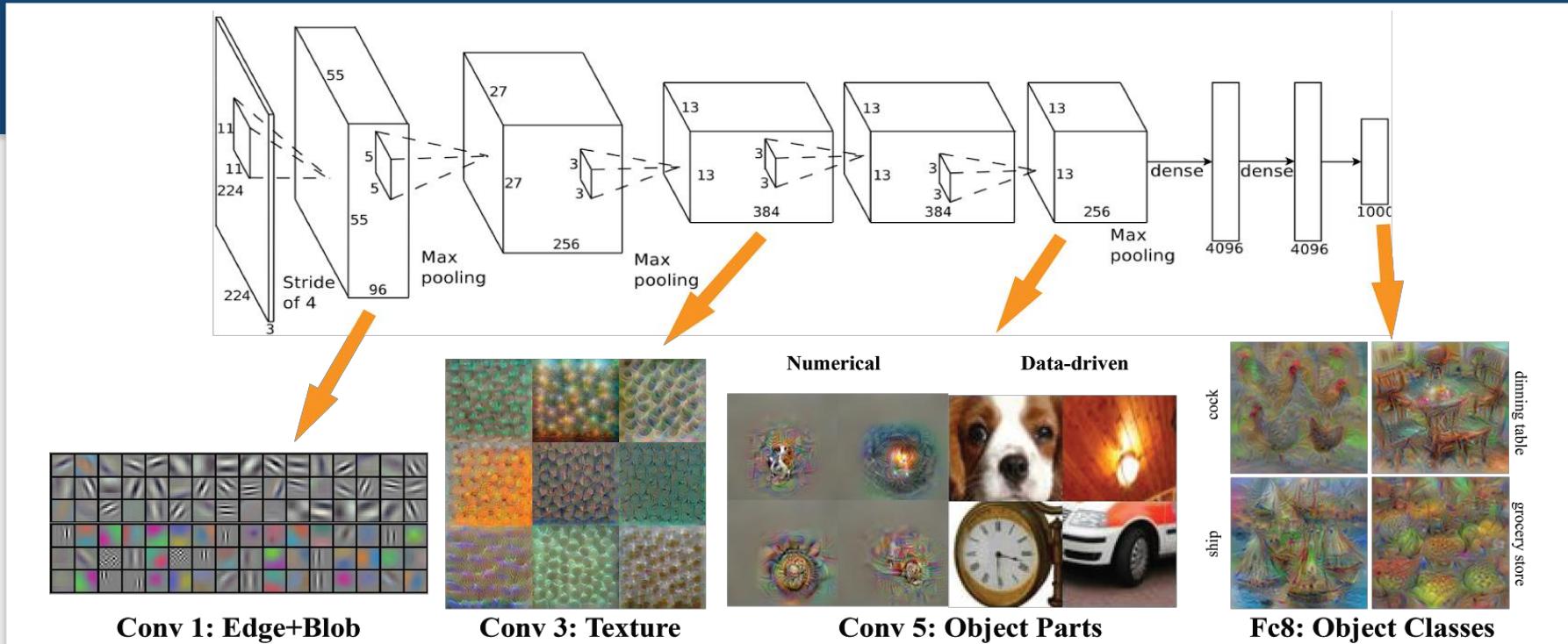
Convolutional Neural Networks learn a complex representation of visual data using vast amounts of data. They are inspired by the human visual system and learn multiple layers of transformations, which are applied on top of each other to extract a progressively more sophisticated representation of the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



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

# Deep Learning - Basics

## Convolutional Neural Nets (CNN)

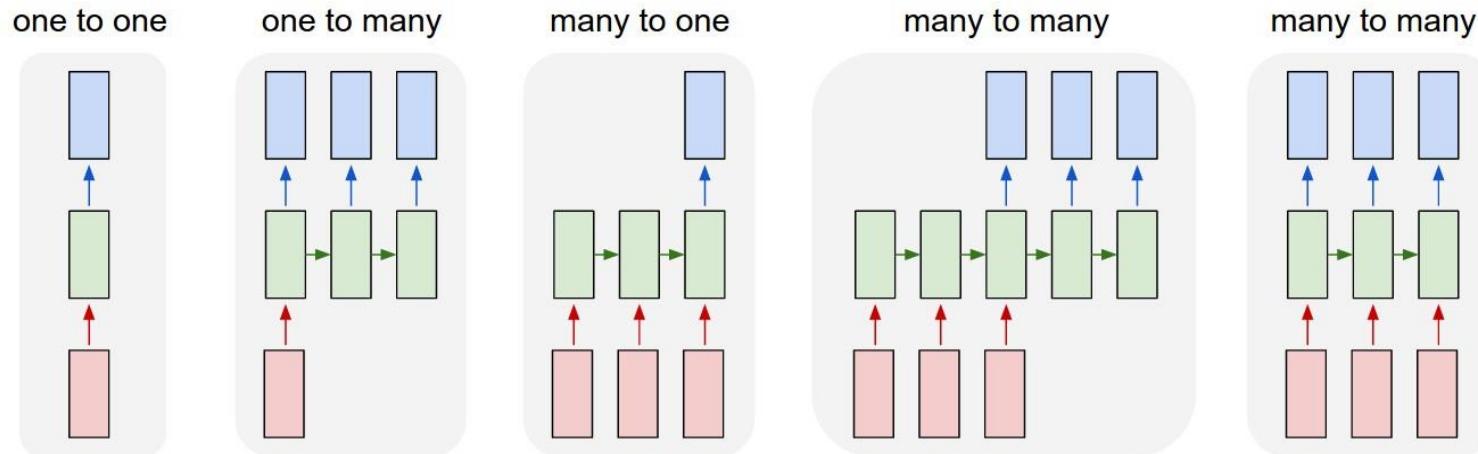


Convolution layer is a feature detector that automagically learns to **filter out not needed information** from an input by using convolution kernel.

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.

# Deep Learning - Basics

## Recurrent Neural Nets (RNN)



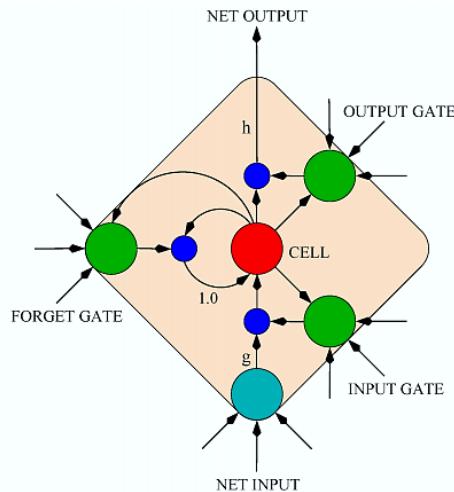
RNNs are general computers which can learn algorithms to map input sequences to output sequences (flexible-sized vectors). The output vector's contents are influenced by the entire history of inputs.



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.*

# 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.



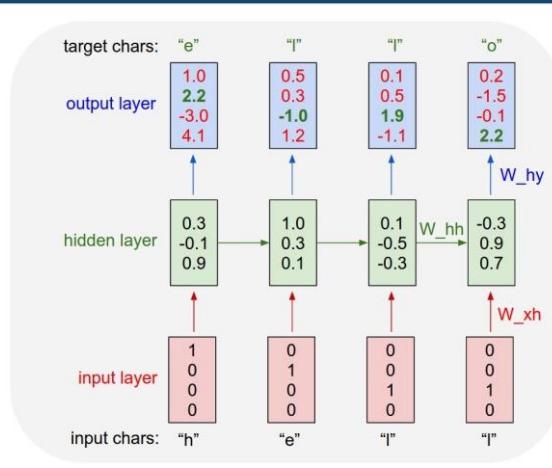
The LSTM units give the network **memory cells with read, write and reset operations**. During training, the network can learn when it should remember data and when it should throw it away.



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**.

# Deep Learning - Basics

## Recurrent Neural Nets (RNN) – Generating Text



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).

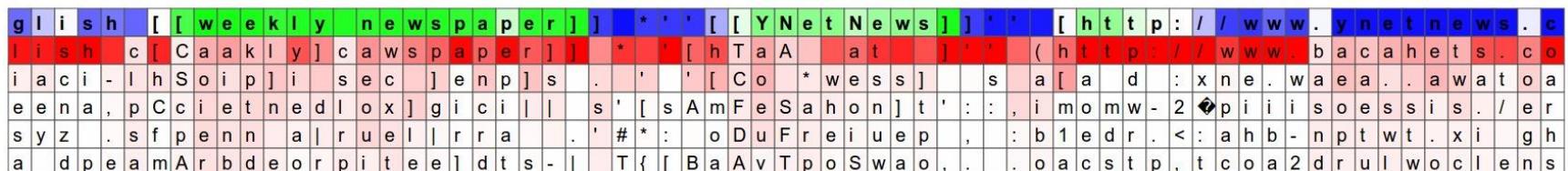
The emperor travelled back to [[Antioch, Perth, October 25|21]] to note, the Kingdom of Costa Rica, unsuccessful fashioned the [[Thrales]], [[Cynth's Dajoard]], known in western [[Scotland]], near Italy to the conquest of India with the conflict. Copyright was the succession of independence in the slop 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.

# Deep Learning - Basics

## Recurrent Neural Nets (RNN) – Generating Text

To **generate text**, we feed a character into the trained RNN and get a distribution over what characters are likely to come next (*red = likely*). We sample from this distribution, and feed it right back in to get the next letter.



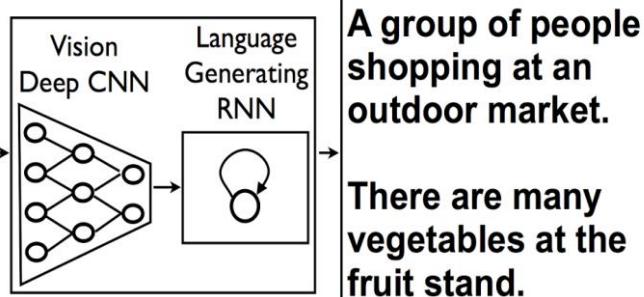
This highlighted neuron gets very excited (*green = excited, blue = not excited*) when the RNN is **inside the [[ ]]** markdown environment and turns off outside of it.



The RNN is likely using this neuron to remember **if it is inside a URL** or not.

# Deep Learning - Basics

## Image Captioning – Combining CNN and RNN



**A group of people shopping at an outdoor market.**  
**There are many vegetables at the fruit stand.**

Neural Image Caption Generator generates fitting natural-language captions only based on the pixels by combining a vision CNN and a language-generating RNN.



A close up of a child holding a stuffed animal



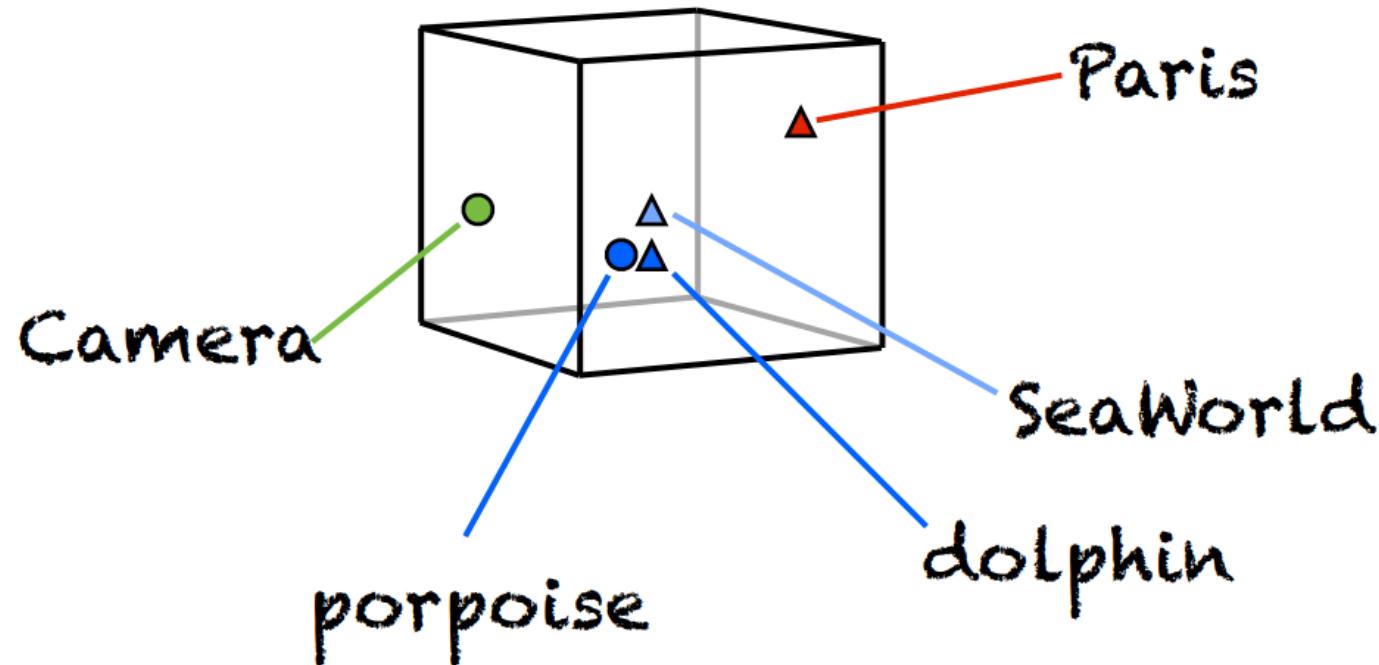
Two pizzas sitting on top of a stove top oven



A man flying through the air while riding a skateboard

# Deep Learning - Basics

## Natural Language Processing – Embeddings

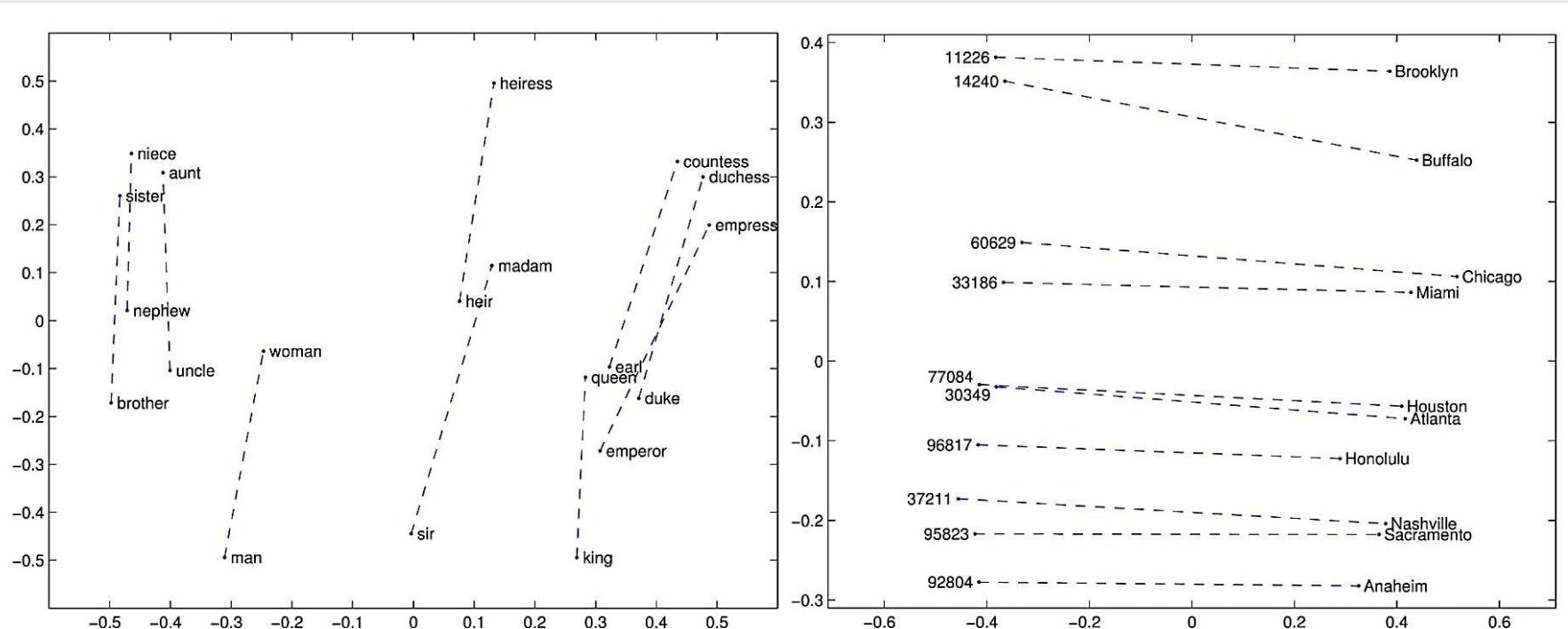


Embeddings are used to turn textual data (words, sentences, paragraphs) into high-dimensional vector representations and group them together with semantically similar data in a **vectorspace**. Thereby, computer can detect similarities mathematically.

# 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.

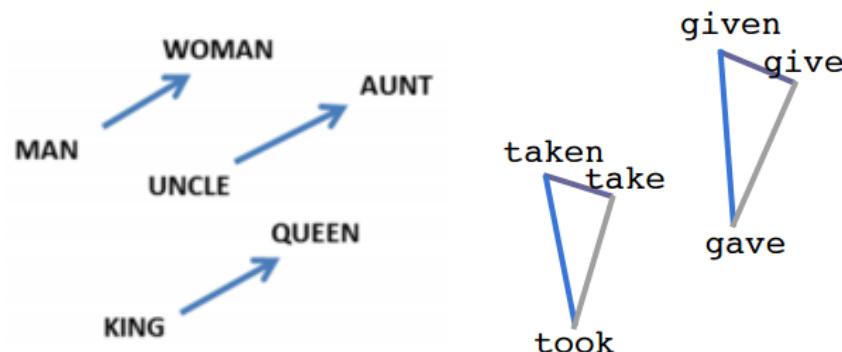


It detects similarities mathematically by grouping the vectors of similar words together. All it needs is **words co-occurrence** in the given corpus.

# 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  $\approx$  Aunt - Uncle  
King - Male + Female  $\approx$  Queen  
Human - Animal  $\approx$  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.

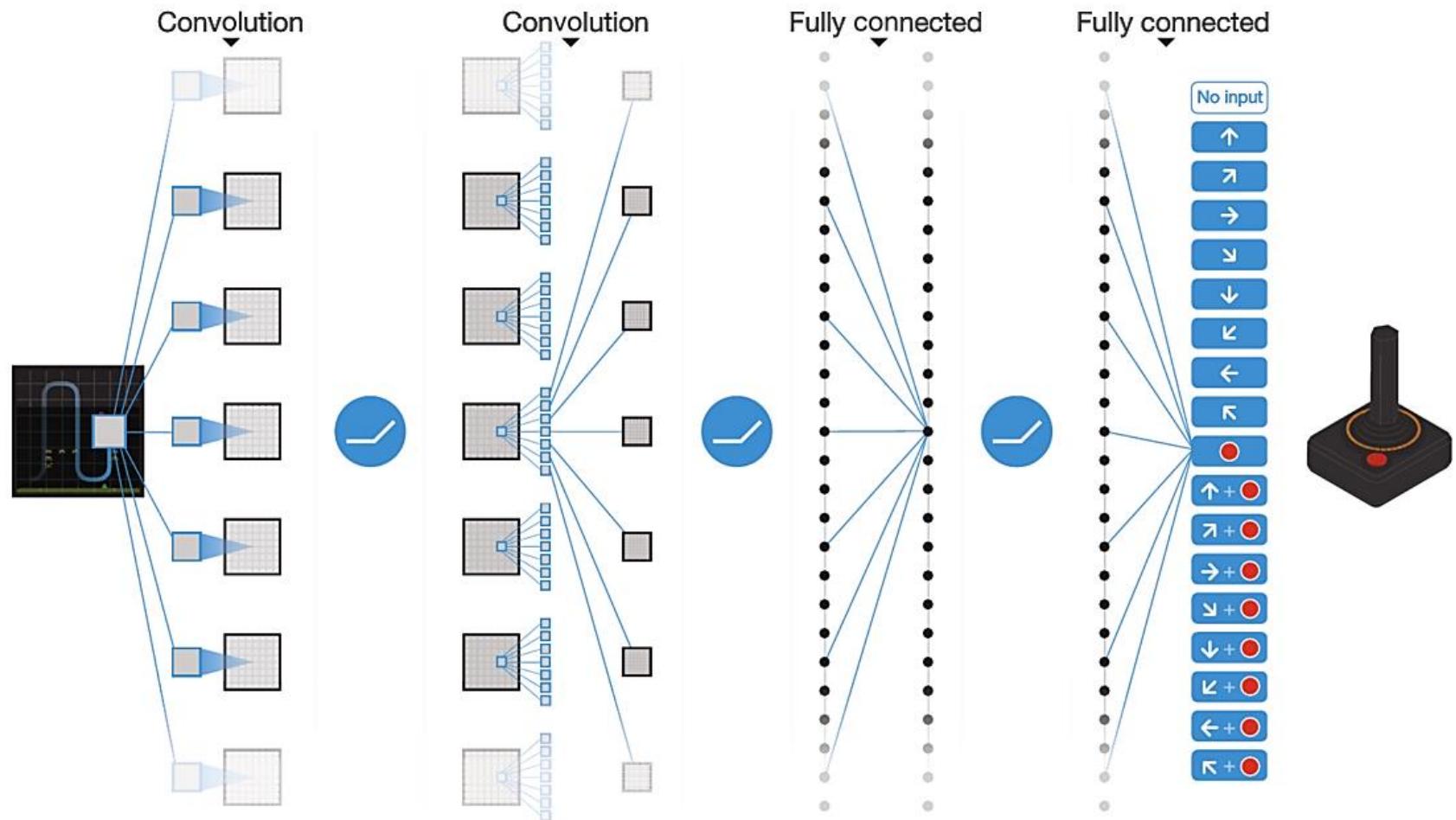
By reading every document on the web, computers might be able to reason like humans do by mimicking the thoughts expressed in content.



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.

# Deep Learning - Basics

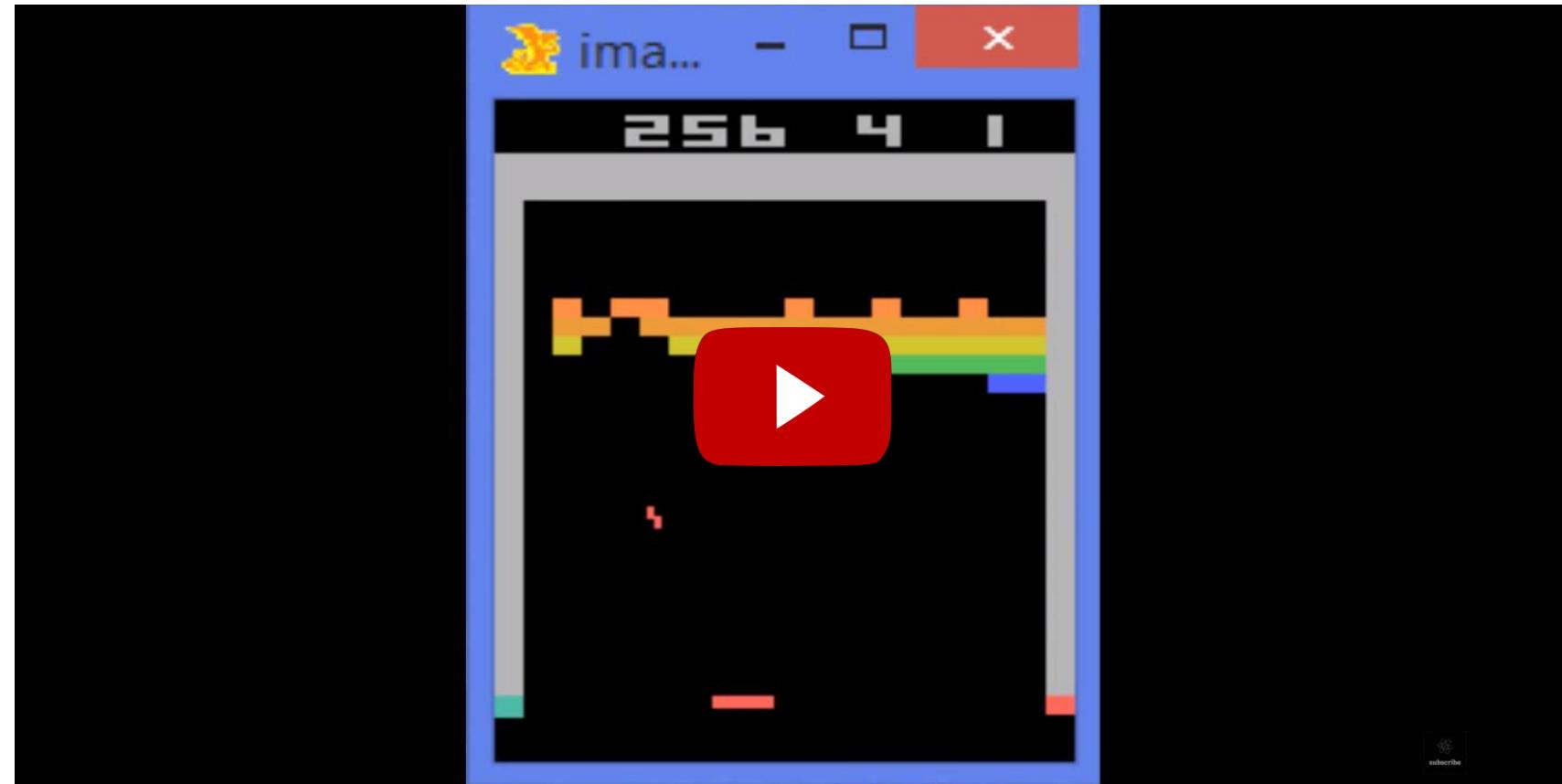
## DeepMind Deep Q-Learning



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

# Deep Learning - Basics

## DeepMind Deep Q-Learning



Outperforms humans in over 30 Atari games just by receiving the pixels on the screen with the goal to maximize the score (Reinforcement Learning)

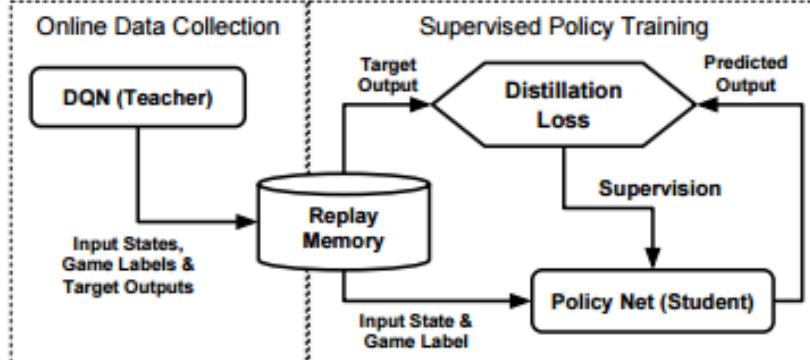
<https://www.youtube.com/watch?v=V1eYniJ0Rnk>

# Deep Learning - Basics

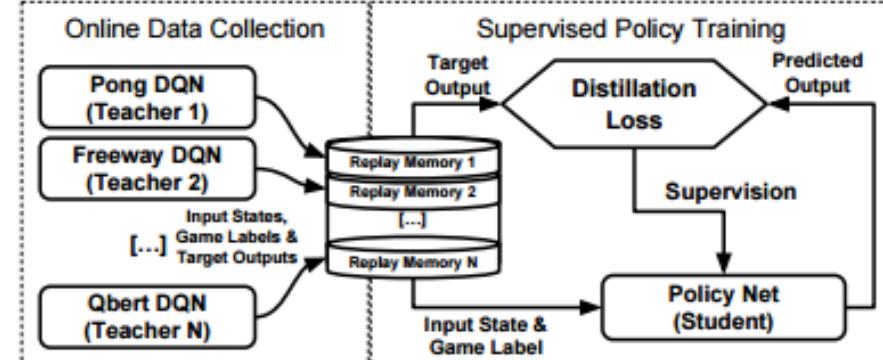
## DeepMind Deep Q-Learning



Policy distillation: Extracts the learned state (*policy*) of a reinforcement learning agent (*teacher*) and **trains a new network (*student*) that performs at the expert level** while being dramatically smaller and more efficient.



Single-task policy distillation



Multi-task policy distillation

# Deep Learning - Basics

## Usage Requirements



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



Measurable and describable goals (*define the cost*)



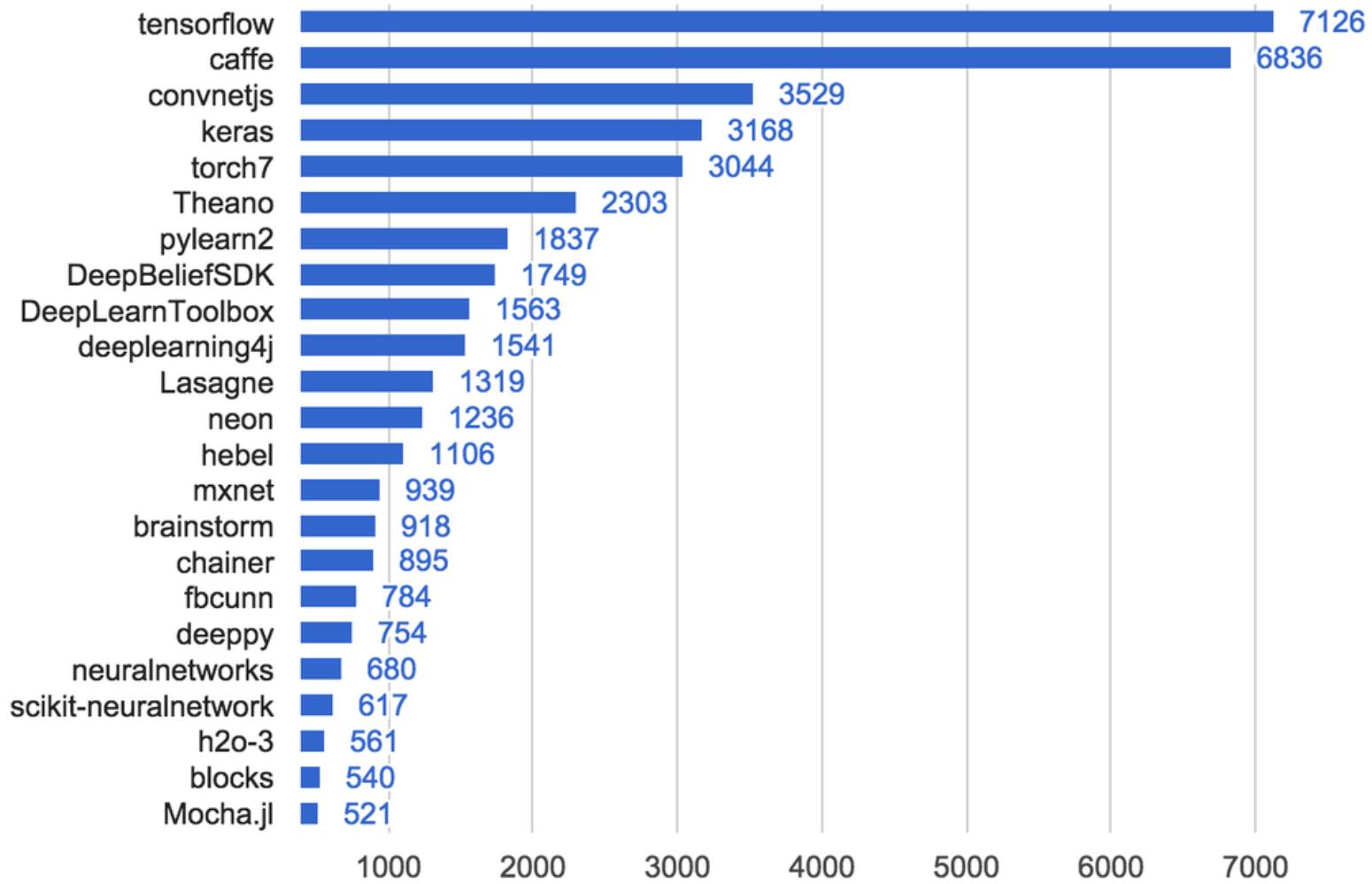
Enough computing power (AWS GPU Instance)



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**

# Deep Learning - Tools

Its all Open Source



# Deep Learning - Tools

Computing is affordable



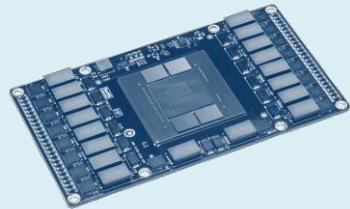
AWS EC2 GPU Spot Instance: *g2.2xlarge* - \$0.0782 per Hour



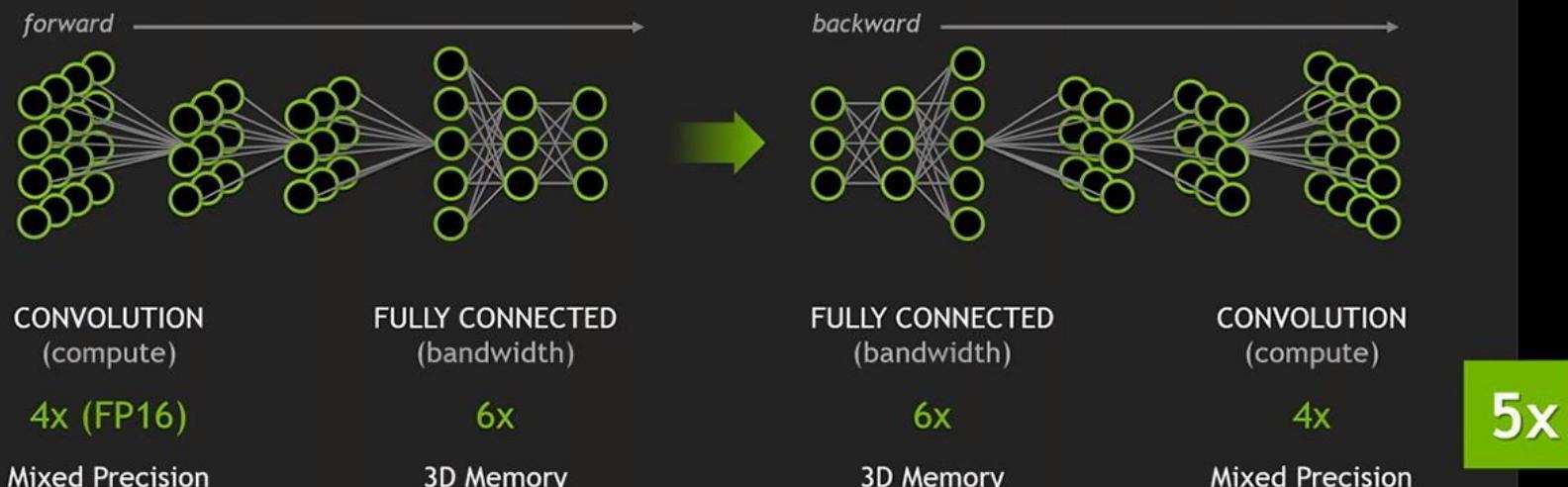
The **DIGITS DevBox** combines the world's best hardware (4 GPUs), software, and systems engineering for deep learning in a powerful solution that can fit under your desk. Cost: \$15k

# Outlook

## NVIDIA Pascal



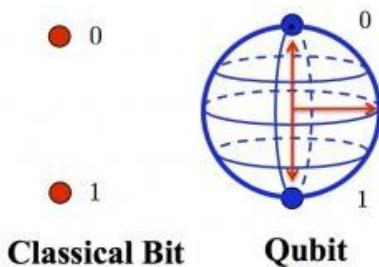
NVIDIA's Pascal GPU architecture will **accelerate deep learning applications** up to 10X beyond the speed of its current-generation Maxwell processors.



# Outlook

## Artificial Quantum Intelligence

Quantum Artificial Intelligence Lab is a joint initiative of NASA and Google to study how quantum computing might advance machine learning. This type of computing may provide the most creative and parallelized problem-solving process under the known laws of physics.



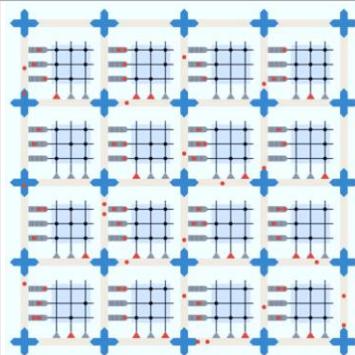
Quantum computers handle what are called quantum bits or qubits that can readily have a value of one or zero or anything in between.



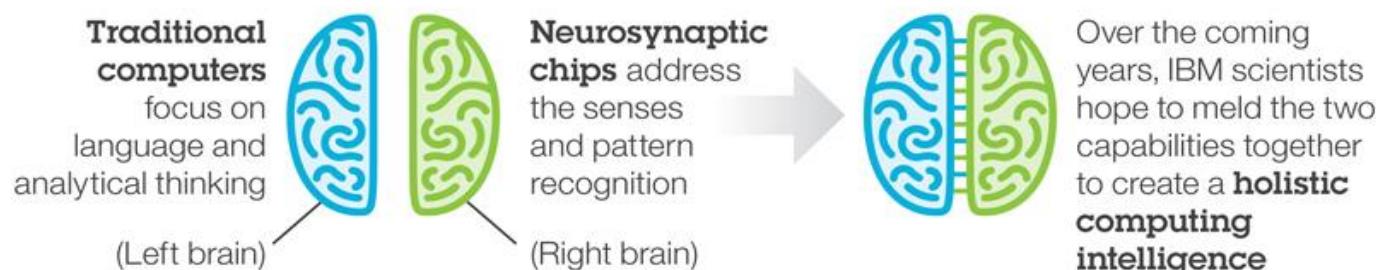
Quantum computing represents a paradigm shift, a radical change in the way we do computing and at a scale that has unimaginable power – *Eric Ladizinsky (Co-founder D-Wave)*

# Outlook

## Neuromorphic Chips



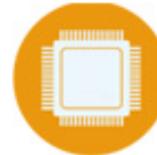
IBM TrueNorth is a brain-inspired computer chip that implements networks of integrate-and-fire spiking artificial neurons and uses only a tiny 70 mw of power –orders of magnitude less energy than traditional chips. The system is designed to be able to run deep-learning algorithms.



1 million  
Programmable  
Neurons



256 million  
Programmable  
Synapses



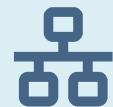
4096  
Neurosynaptic  
Cores

# Outlook

## Deep Learning



Significant advances in **deep reinforcement and unsupervised learning**



Bigger and **more complex architectures** based on various interchangeable modules/techniques



**Deeper models** that can learn from much fewer training cases



Harder problems such as **video understanding** and **natural language processing** will be successfully tackled by deep learning algorithms

# Takeaways



Machines that **learn to represent the world** from experience.



Deep Learning is **no magic!** Just statistics in a black box, but exceptional effective at learning patterns.



We haven't figured out **creativity** and **human-empathy**.



Transitioning from research to consumer products. Will make the tools you use every day **work better, faster and smarter**.