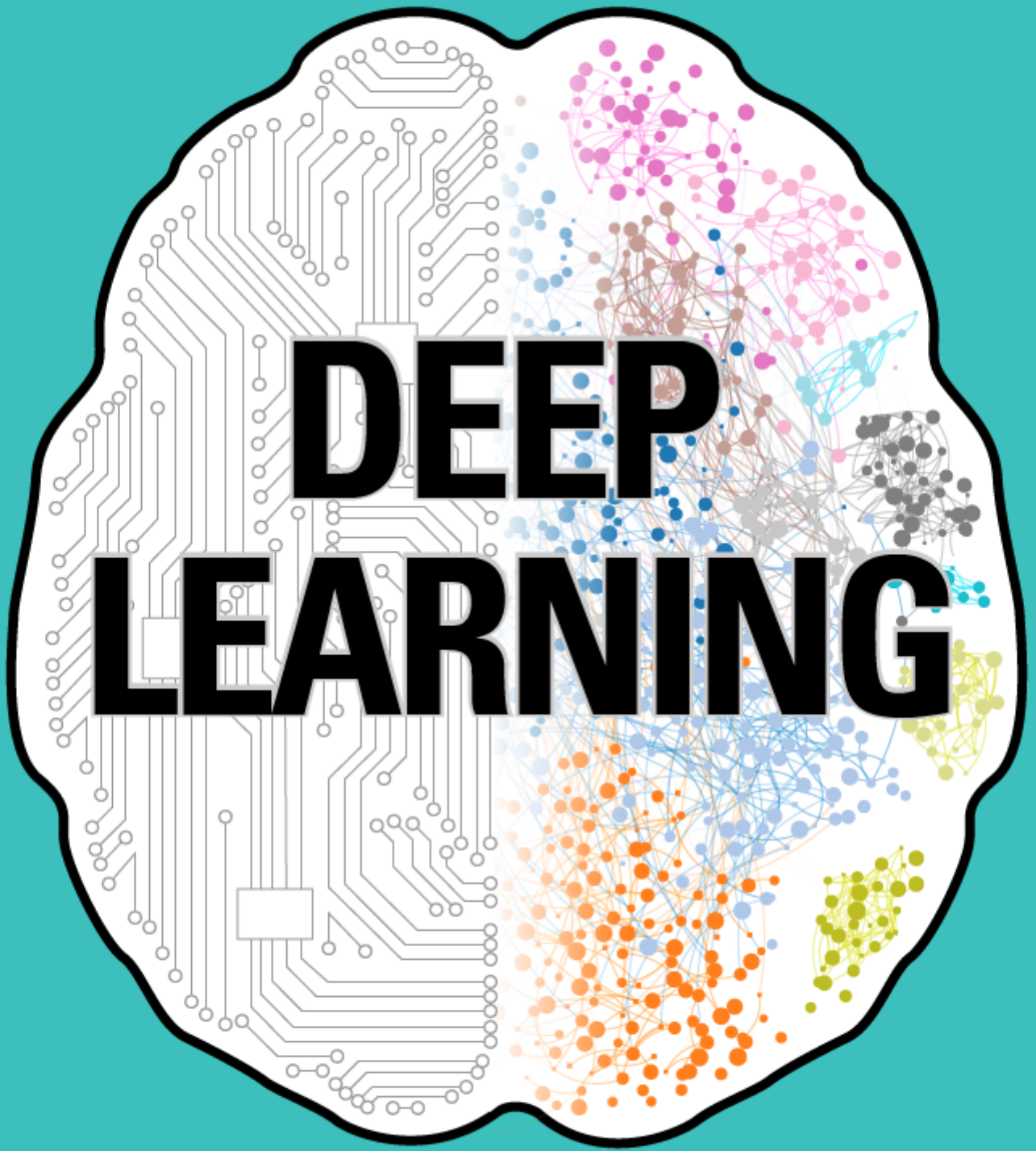

DEEP
LEARNING

Presented By:

Sourav Kumar Singh
Gopal Krishna
Subhankar Halder
Sayan Khan

Contents in Brief


Introduction

What is learning?

The *acquisition of knowledge* or skills through study, experience, or being taught.

What is AI?

Intelligence: ”the ability to learn, understand and think”(oxford dictionary)

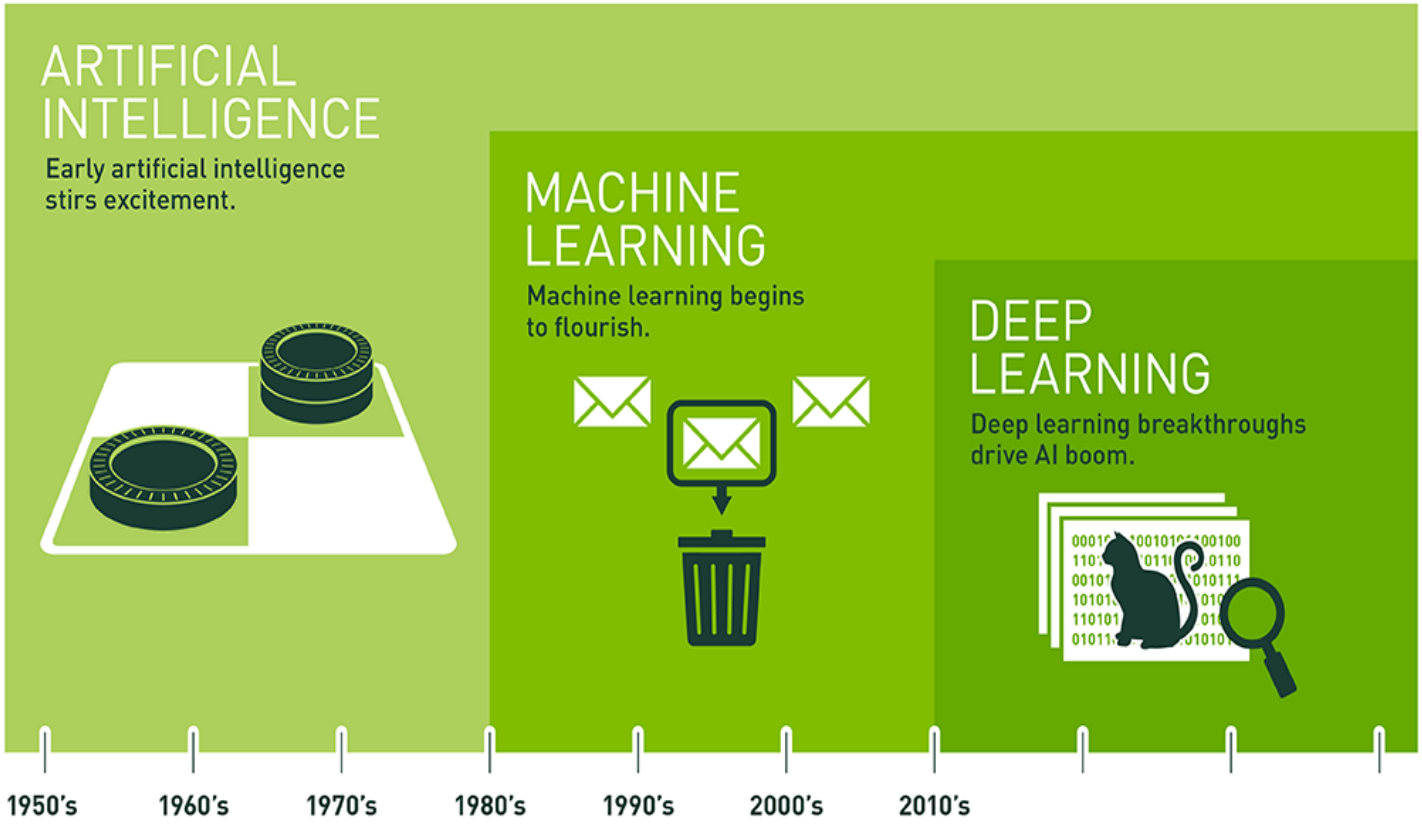
Definition

AI is the study of how to make computers do things which at the moment humans do better.

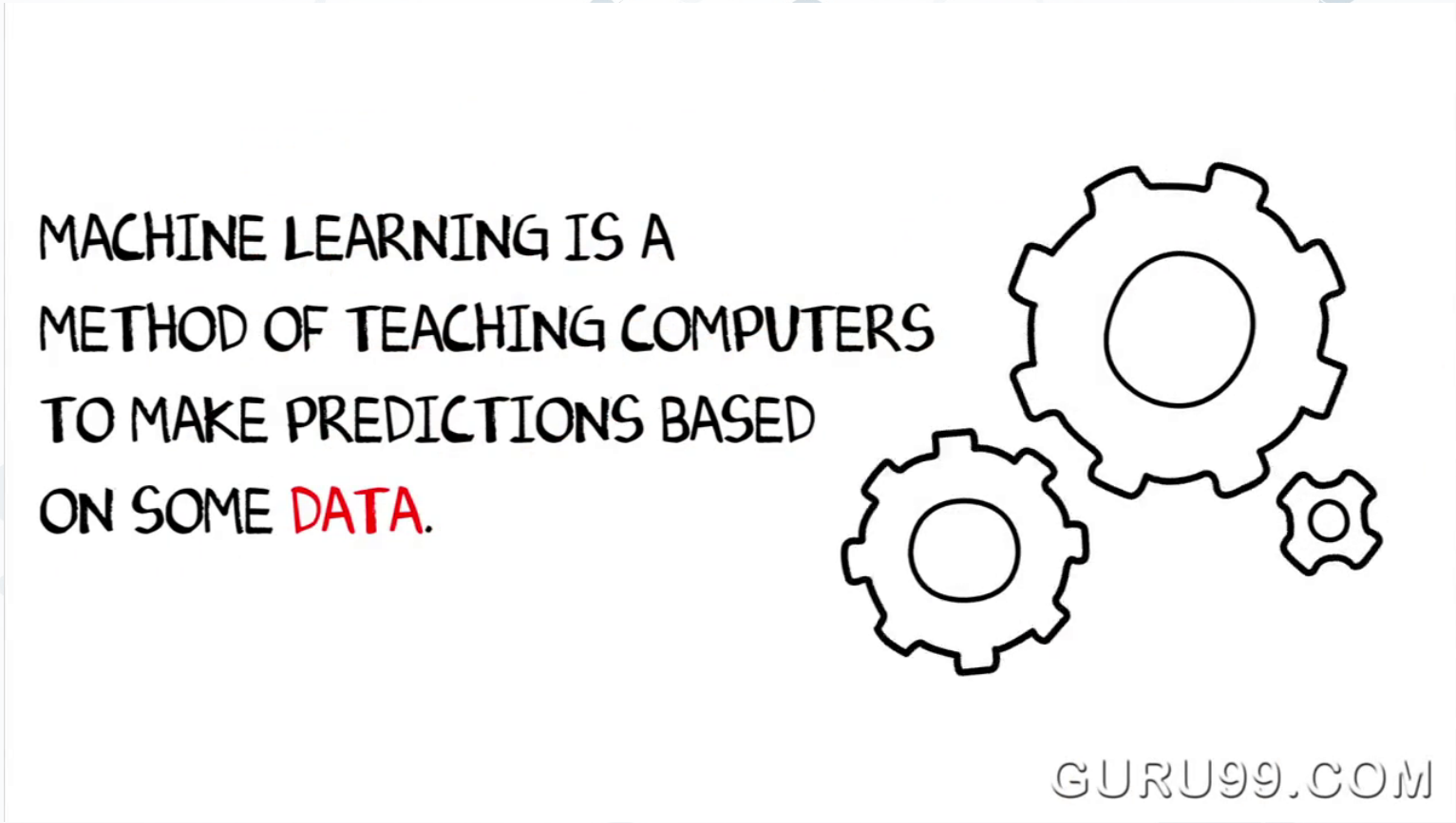
Examples include Speech Recognition; Smell, Face, Object Detection; Intuition and Inferencing; Decision making; Abstract Thinking

The easiest way to think of their relationship is to visualize them as concentric circles with AI - the idea that came first - the largest, then machine learning - which blossomed later, and finally deep learning - which is driving todays AI explosion - fitting inside both.

Deep Learning is an extension of AI



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.



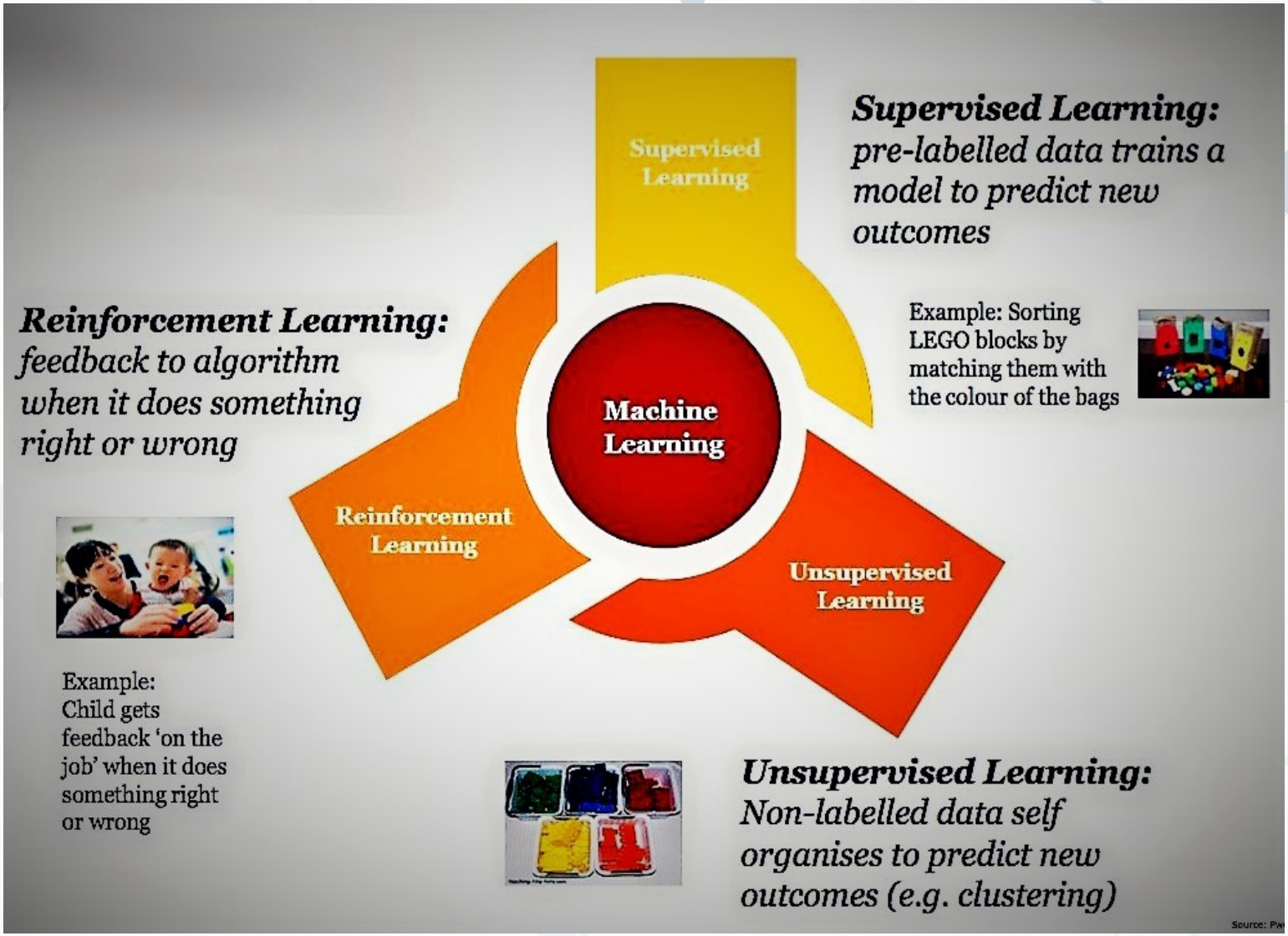
Definition

Machine learning is the science of getting computers to act without being explicitly programmed.

In the past decade, machine learning has given us self-driving cars, practical speech recognition, effective web search, and a vastly improved understanding of the human genome. Machine learning is so pervasive today that you probably use it dozens of times a day without knowing it. Many researchers also think it is the best way to make progress towards human-level AI.

Machine learning is needed for tasks that are too complex for humans to code directly.

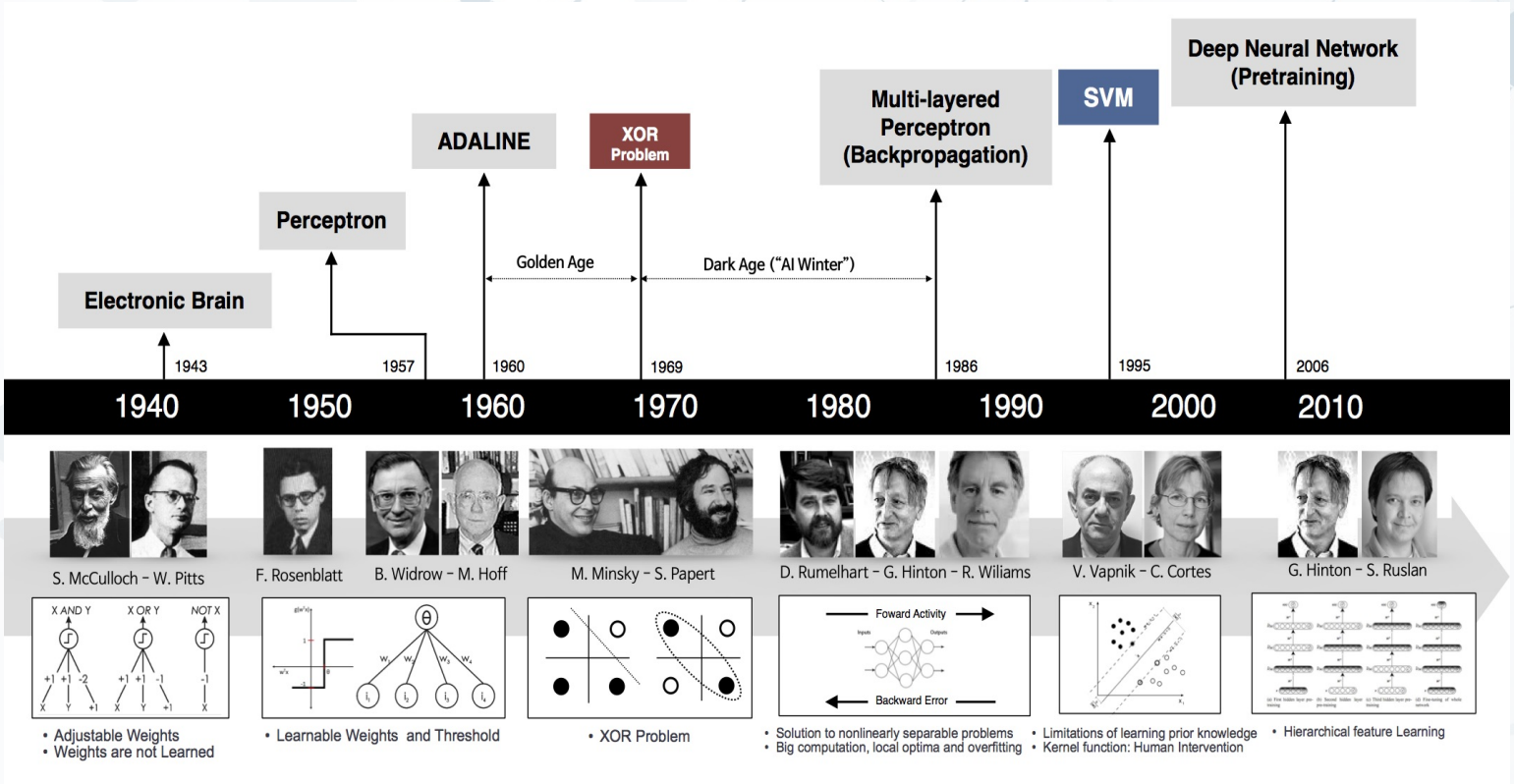

Data mining
Algorithms
MACHINE LEARNING
Strategy
Artificial intelligence
Statistics
Prediction
Analysis


Supervised Learning
Machine Learning
Reinforcement Learning
Unsupervised Learning
Supervised Learning: pre-labelled data trains a model to predict new outcomes
Example: Sorting LEGO blocks by matching them with the colour of the bags
Reinforcement Learning: feedback to algorithm when it does something right or wrong
Example: Child gets feedback 'on the job' when it does something right or wrong
Unsupervised Learning: Non-labelled data self organises to predict new outcomes (e.g. clustering)

Common Classes of Algorithms

(Supervised | Unsupervised)

A little look into the history



- **John McCarthy** coined the term *"artificial intelligence"* as the topic of the Dartmouth Conference, the first conference devoted to the subject.
- **Arthur Samuels** is recognised as the first learning machine which learned to play checkers. His algorithms used a heuristic search memory to learn from experience.
- 1970 - 1980 - **AI WINTER**
- In mid 1980s, Backpropagation algorithm is discovered which allowed more powerful neural networks with more hidden layers to be trained.

- By mid 1990s, new methods like SVMs become in-vogue and are considered amenable to rigorous mathematical analysis and achieve state-of- the art performance.
- At the same time, major advances in all areas of AI, with significant demonstrations in machine learning, intelligent tutoring, data mining, natural language understanding and translation, vision, virtual reality, games, and other topics take place.
- 2006 - Rebranding of neural networks to deep learning.
- 2012 - Relative break through using deep nets for speech recognition where, for one of the first times, a neural model achieved state of the art.


DEEP
LEARNING

Robust

- Works on raw data (pixels, sound, text or chars), no need for feature engineering
- Robustness to natural variations in data is automatically learned

Robust

- Works on raw data (pixels, sound, text or chars), no need for feature engineering
- Robustness to natural variations in data is automatically learned

Generalizable

- Allows end-to-end learning (pixels-to-category, sound to sentence, English sentence to Chinese sentence, etc)
  - No need to do segmentation etc. (a lot of manual labour)

Robust

- Works on raw data (pixels, sound, text or chars), no need for feature engineering
- Robustness to natural variations in data is automatically learned

Generalizable

- Allows end-to-end learning (pixels-to-category, sound to sentence, English sentence to Chinese sentence, etc)
  - No need to do segmentation etc. (a lot of manual labour)

Scalable

- Performance increases with more data, therefore method is massively parallelizable

Robust

- Works on raw data (pixels, sound, text or chars), no need for feature engineering
- Robustness to natural variations in data is automatically learned

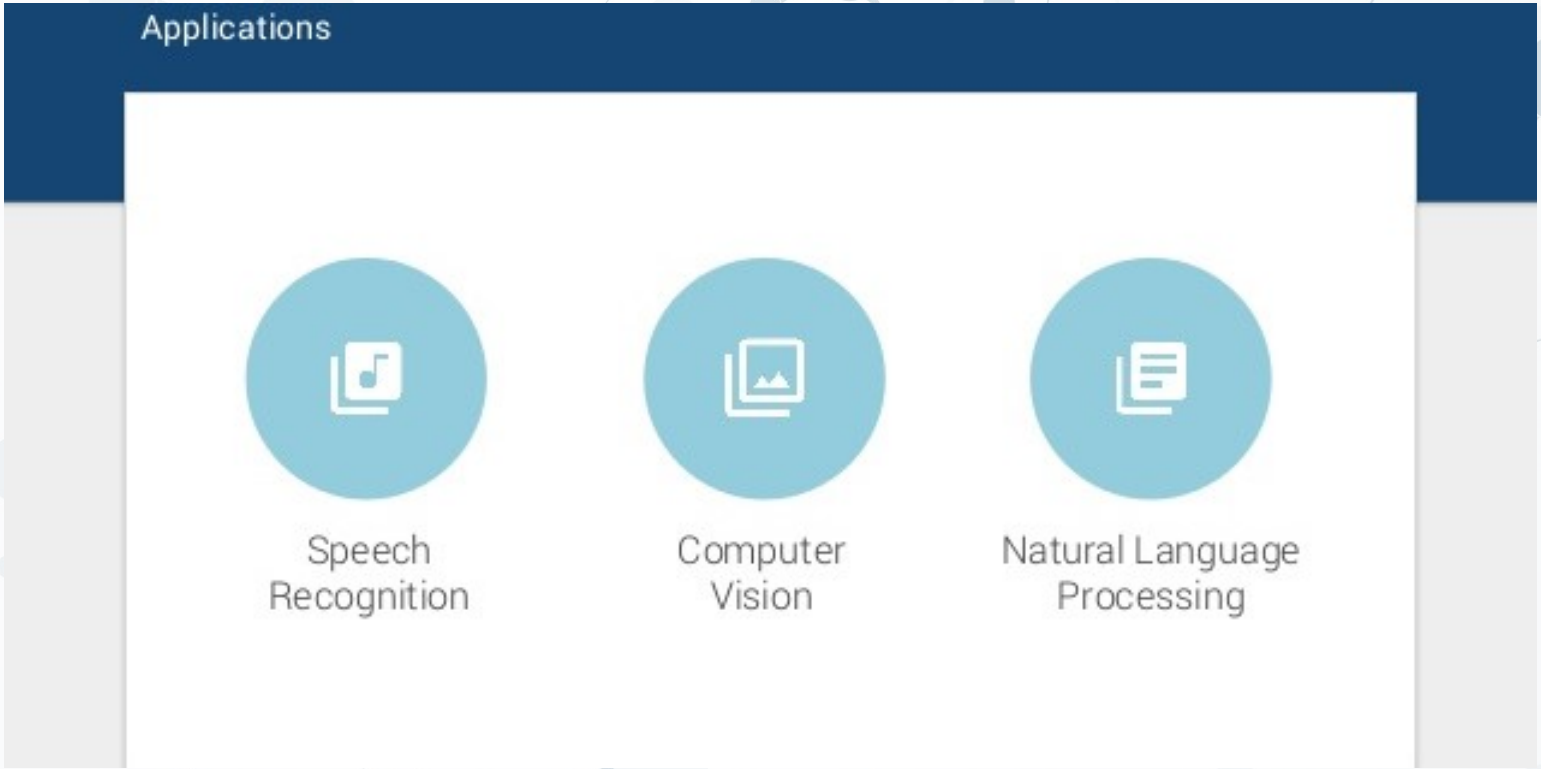
Generalizable

- Allows end-to-end learning (pixels-to-category, sound to sentence, English sentence to Chinese sentence, etc)
  - No need to do segmentation etc. (a lot of manual labour)

Scalable

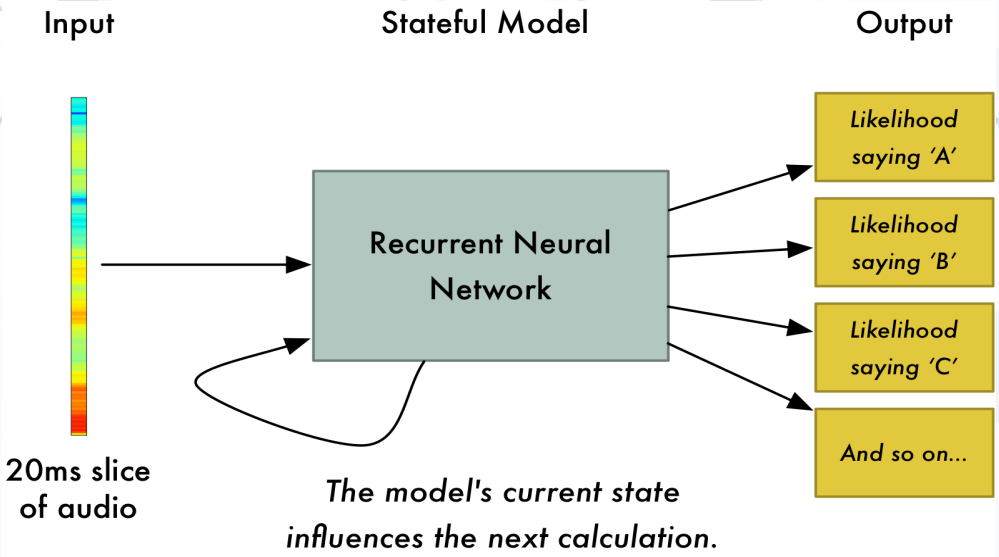
- Performance increases with more data, therefore method is massively parallelizable

You can iterate faster (and get superior quality at the same time!)



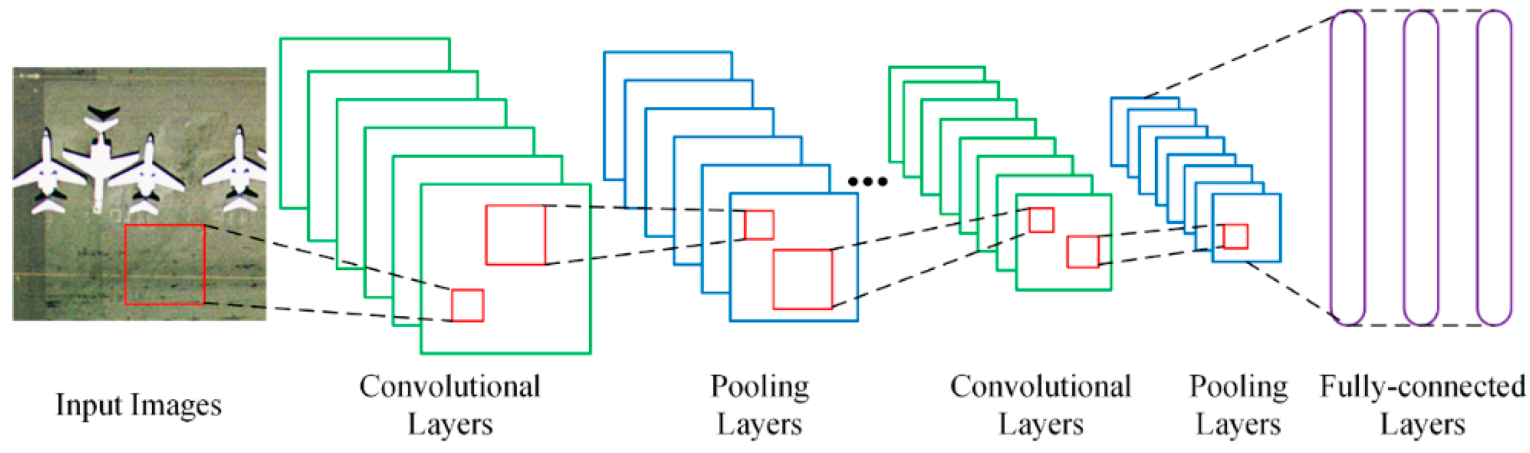
Deep Learning in Speech Recognition

- Deep learning finally made speech recognition accurate enough to be useful outside of carefully controlled environments.
- Andrew Ng has long predicted that as speech recognition goes from 95% accurate to 99% accurate, it will become a primary way that we interact with computers.
- In a format thats easy to process, we will feed it into a deep neural network. For each little audio slice, it will try to figure out the letter that corresponds the sound currently being spoken.



Deep Learning in Computer Vision

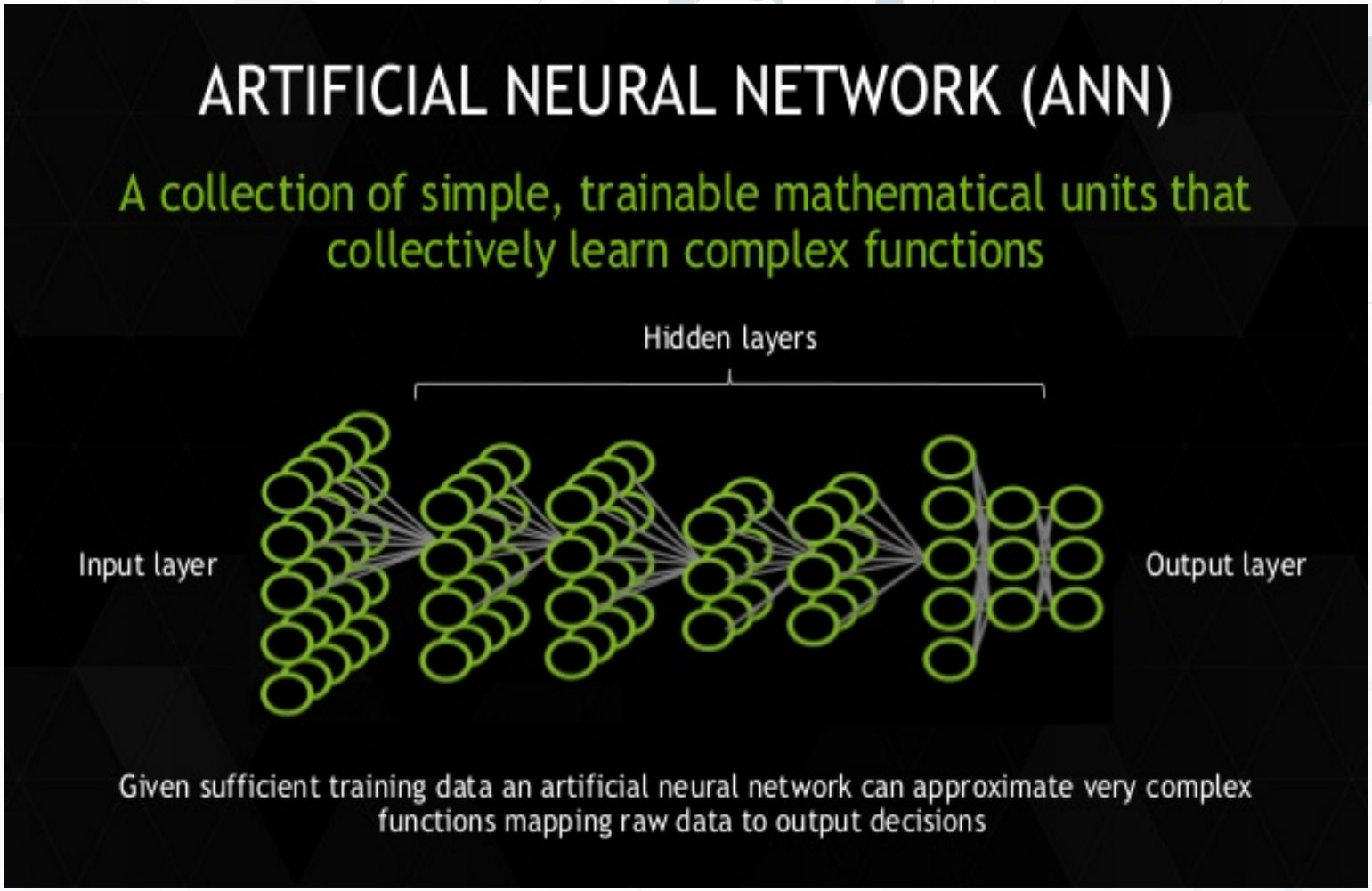
- Learning of local features , like other classical feature detectors like SIFT or SURF, but at different levels and different layers.
- Detectors that learn, they detect different things, different properties of the image at different levels.
- And as we keep moving on, at the final layers we come up with detectors that are even more complicated. Like, they might react to torsos or faces.



Deep Learning in NLP

- Several big improvements in recent years in NLP with different
  - Levels - speech, words, syntax, semantics
  - Tools - parts of- speech, entities, parsing
  - Applications machine translation, sentiment analysis, dialogue agents, question answering
- Neural networks can accurately determine the structure of sentences, supporting interpretation.
  Representaion of NLP Levels for Semantics:
  - Every word and every phase and every logical expression is a vector.
  - A neural network combines two vectors into one vector.

Let's see-



Important Property of Neural Networks

Results get better with

more data +
bigger models +
more computation

(Better algorithms, new insights and improved techniques always help, too!)

Definition

Deep learning is the subfield of Machine Learning which is concerned with learning algorithms that derive meaning out of data by using a ***hierarchy of multiple layers*** that mimic the neural networks of our brain.

These learning algorithms are what we call **Artificial Neural Networks**. They're a special class of algorithms that learn and improve on their own.

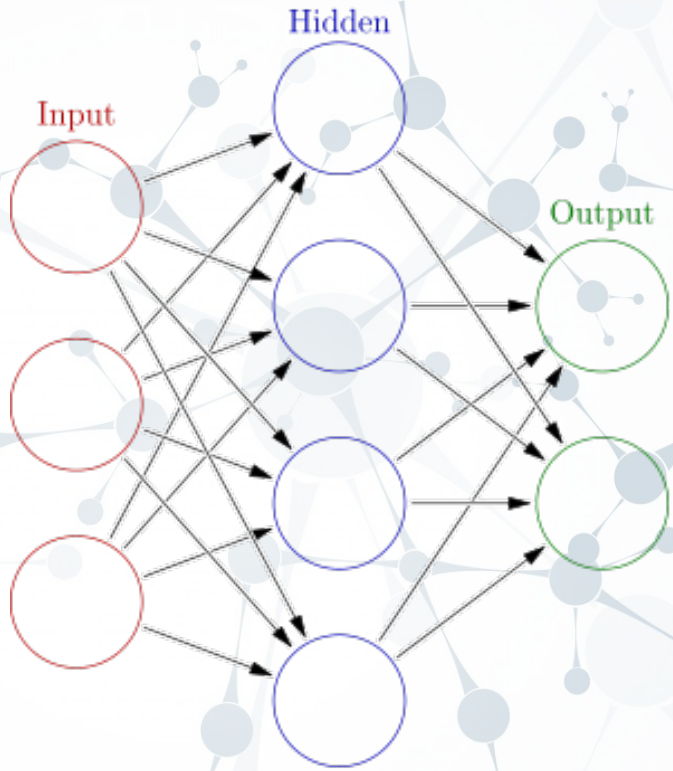


Figure: Hierarchy in simplest kind of Neural Network.

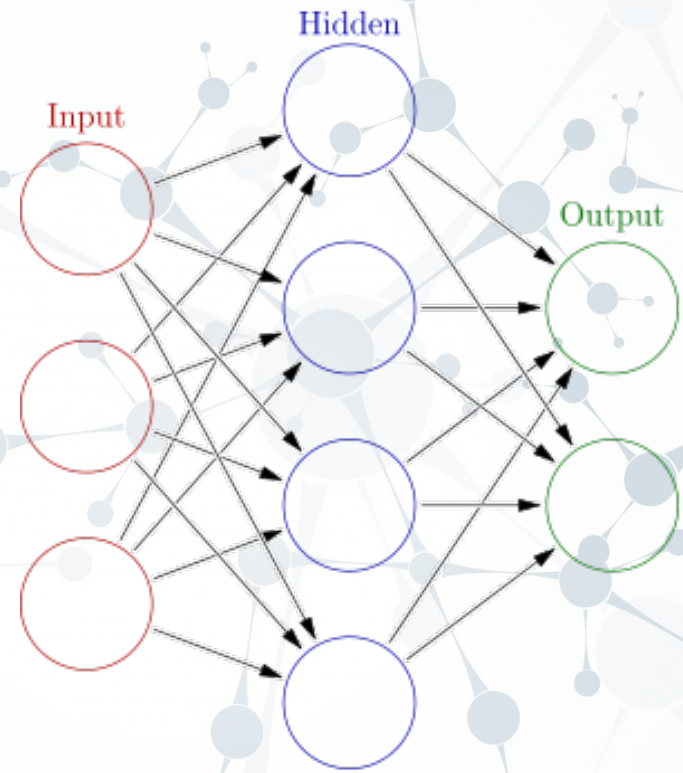


Figure: Hierarchy in simplest kind of Neural Network.

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

So what actually is Deep Learning?

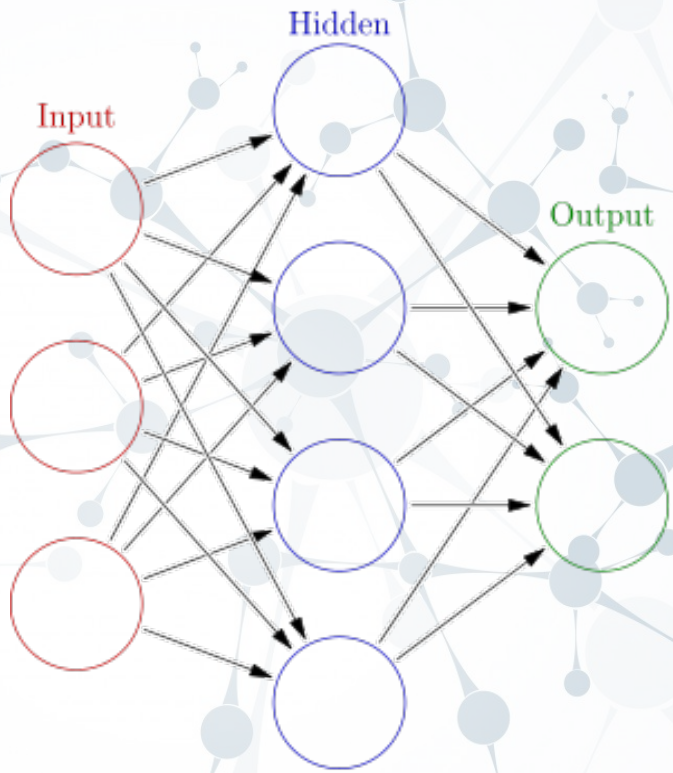
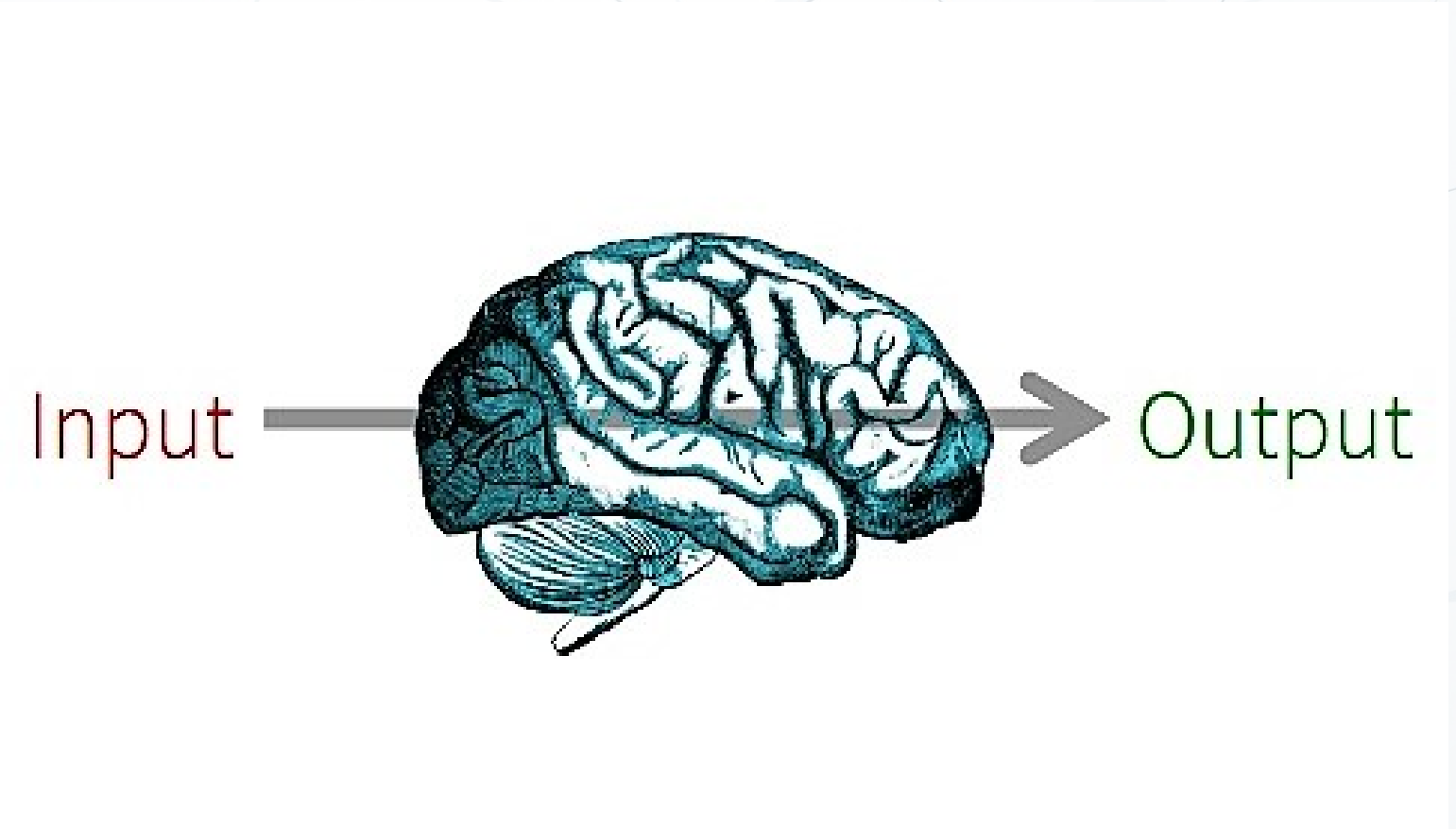
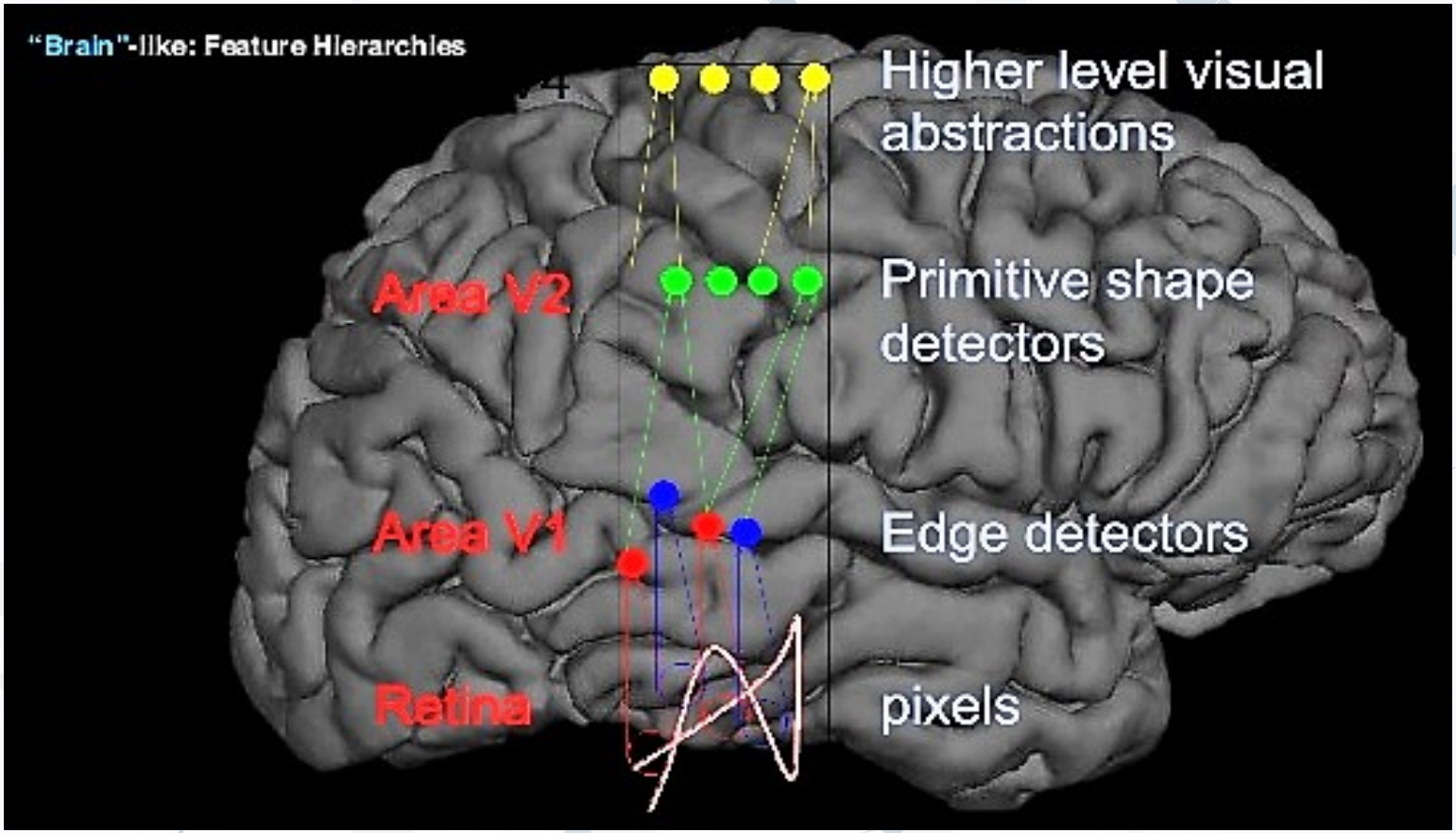


Figure: Hierarchy in simplest kind of Neural Network.

If you provide the system tons of information, it begins to understand it and respond in useful ways. Amazing! Isn't it?

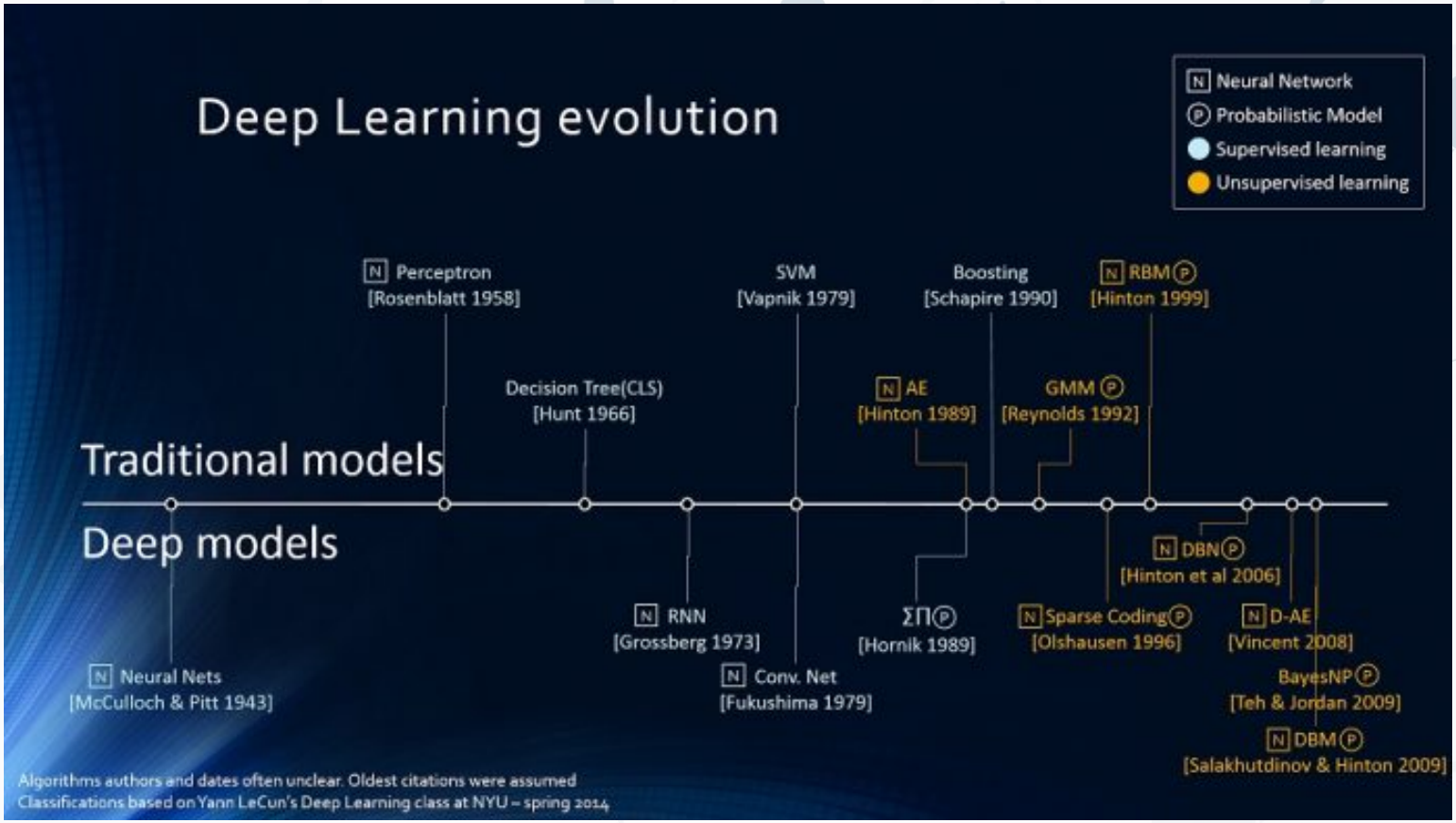
Our brain has a lot of neurons connected together and the strength of the connections between neurons represents long term knowledge.

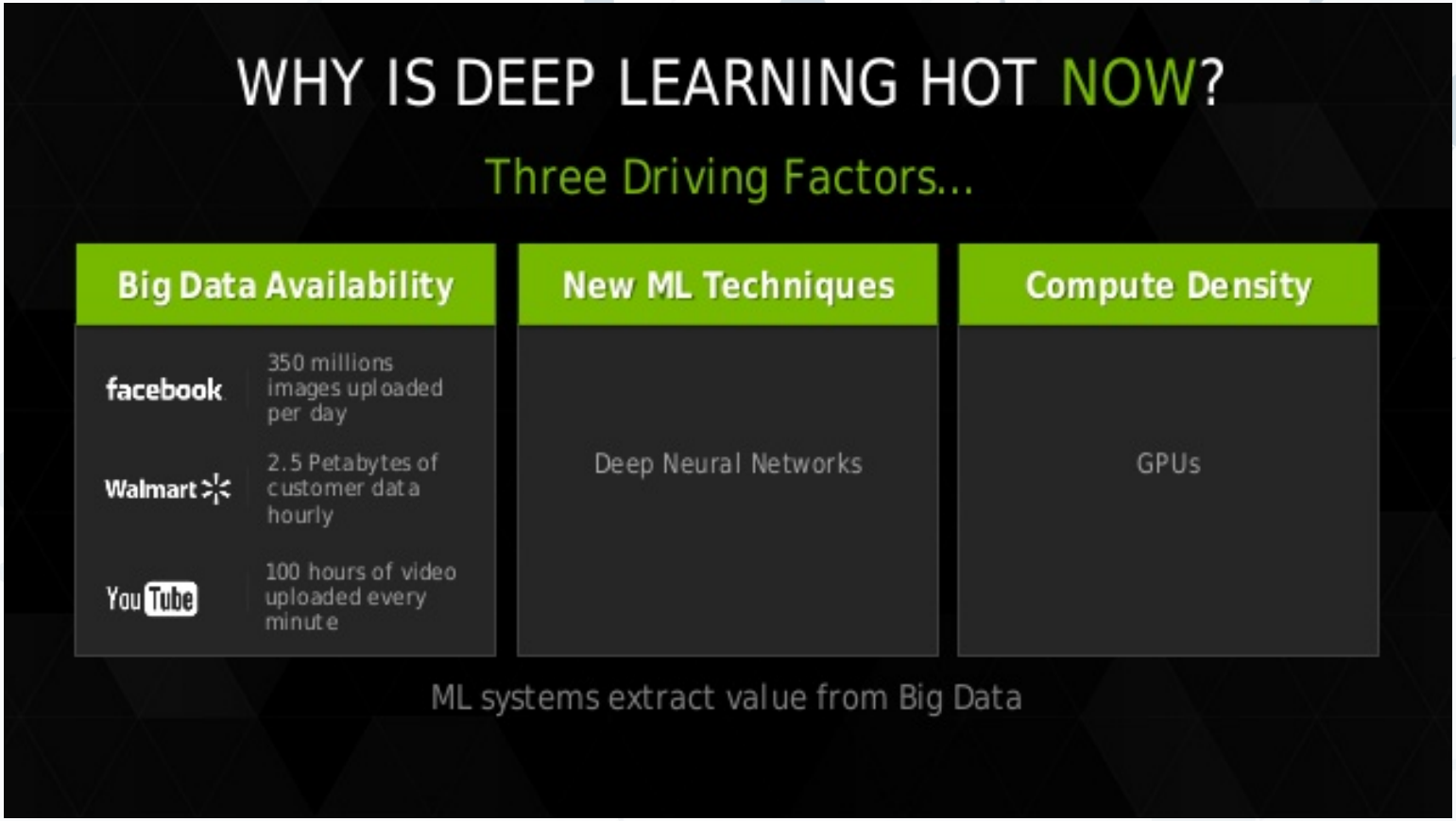




The first hierarchy of neurons i.e V1 that recieves information in the visual cortex are sensitive to specific edges, while the brain regions further down the visual pipeline i.e V2 are sensitive to more complex shapes such as lips, noses etc.

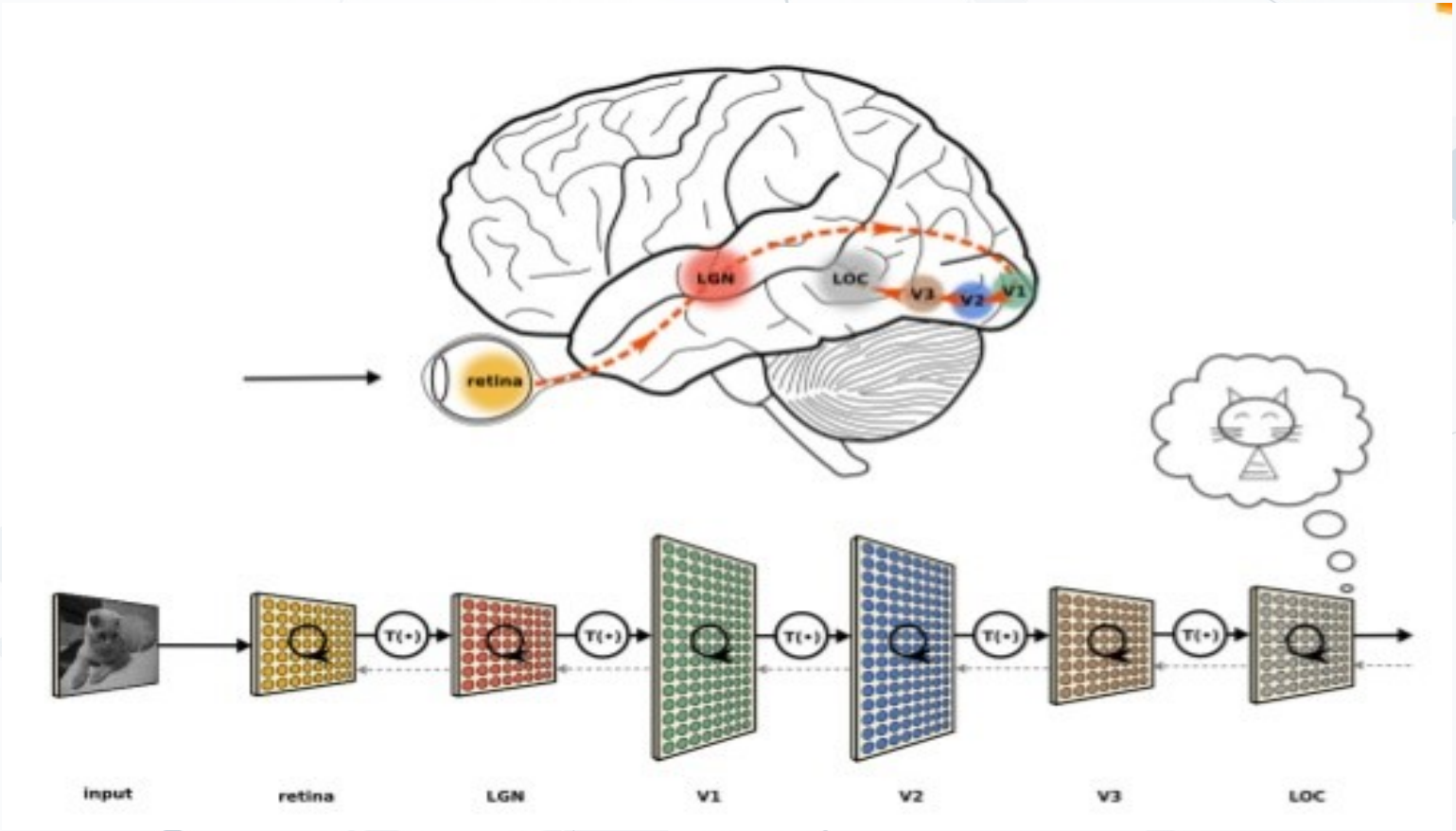
So how did it evolve?





Deep Learning - Basics

How does it work?

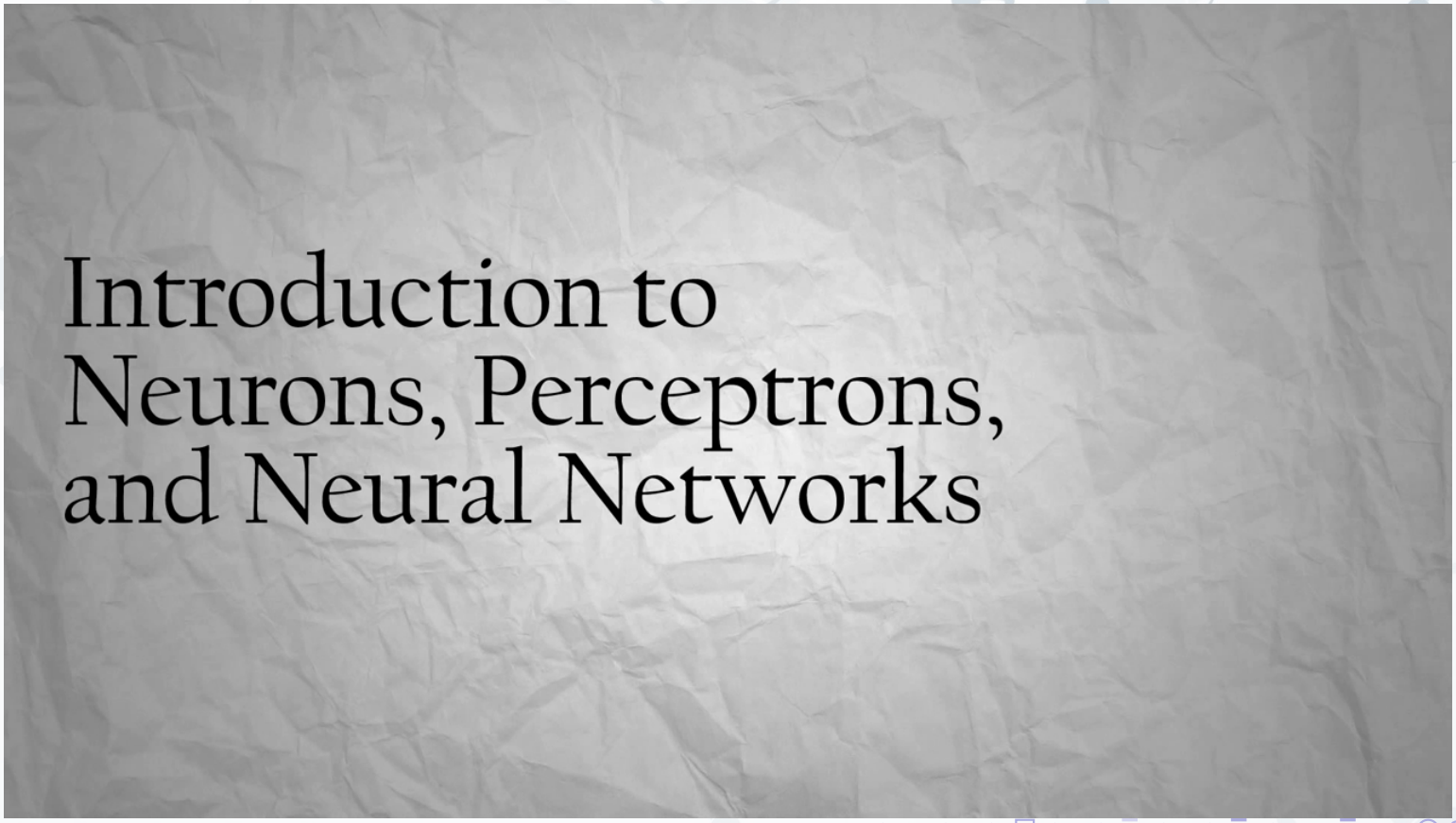


A deep neural network consists of various layers, whereby each layer ***transforms the input data*** into more abstract representations (e.g edge - nose - face). The output layer combines those features to make predictions.

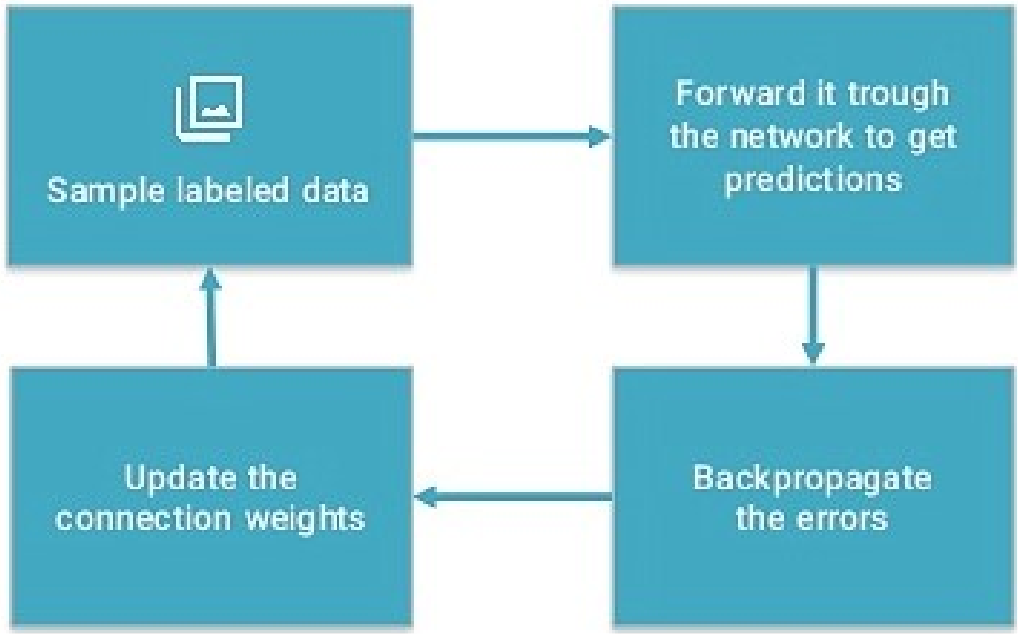
Definition

In the context of neural networks, a **perceptron** is an artificial neuron i.e. the most basic form of an artificial neural network.

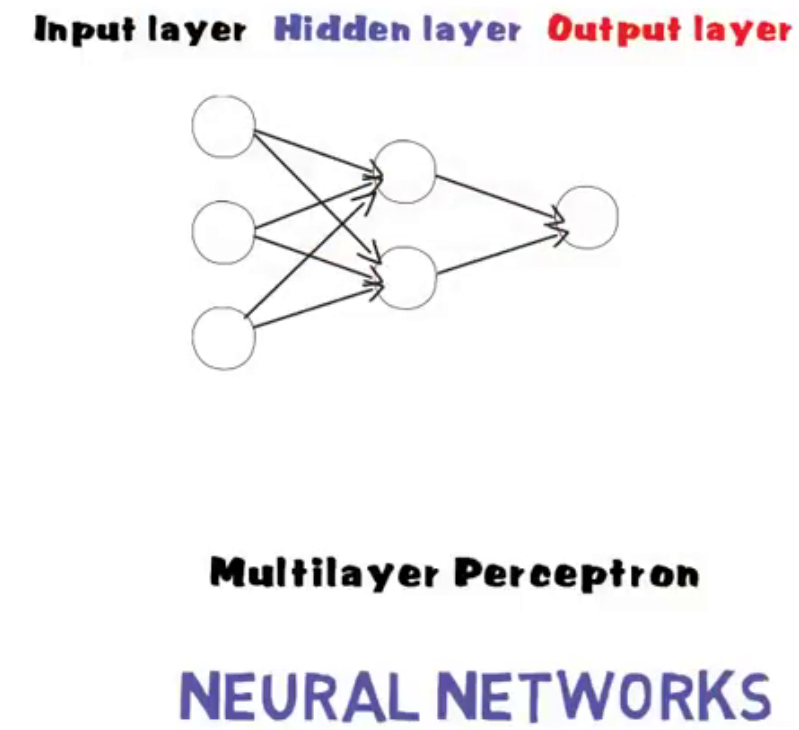
neuron : neural network :: perceptron : ANN



The Training Process



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



- Large data set with good quality (input-output mappings)
- Measurable and describable goals (define the cost)
- Enough computing power (AWS-GPU instance)
- Excels in task where the basic unit (pixels, words) has very little meaning in itself, but the combination of such units has a useful meaning.

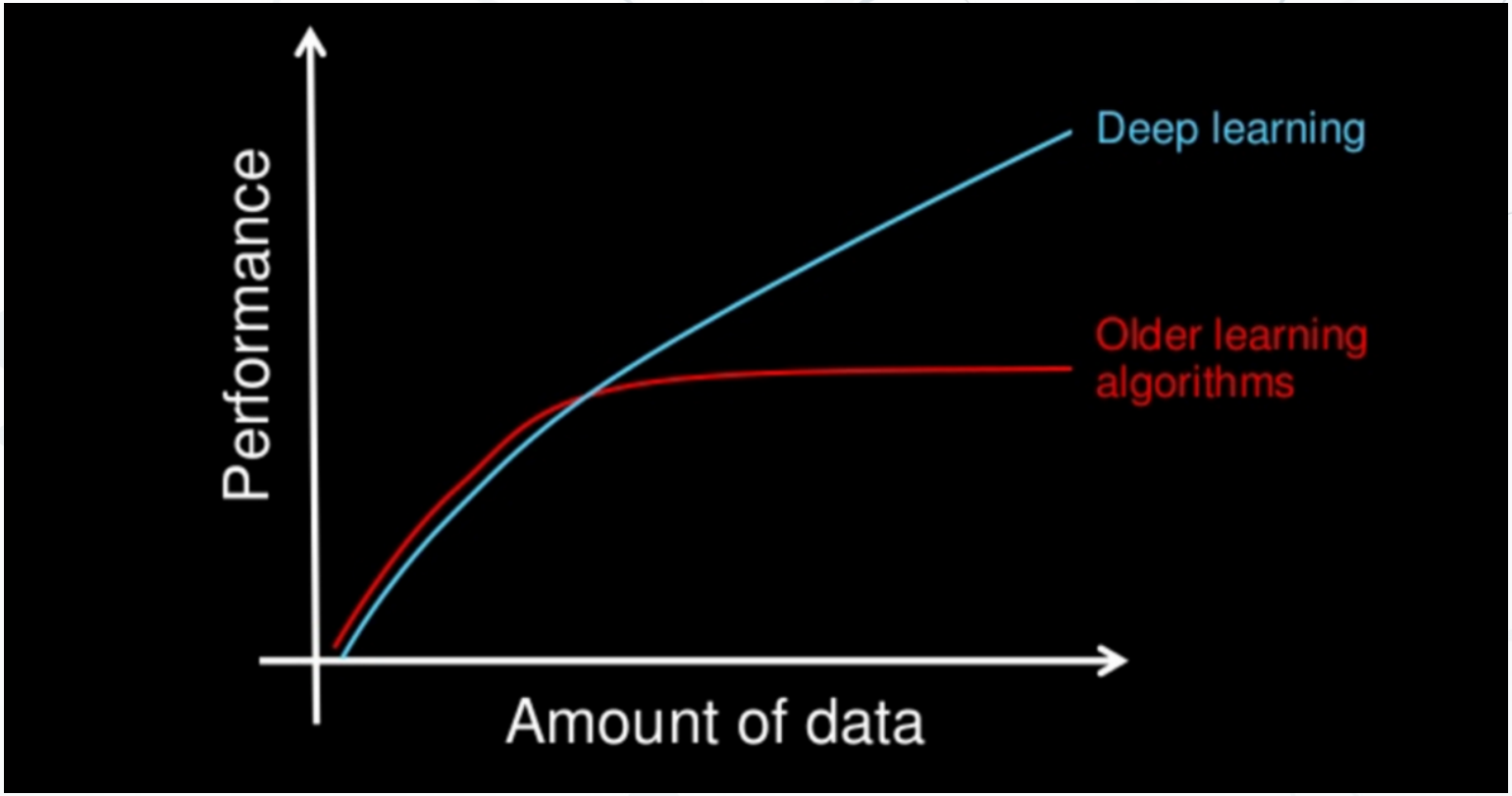
The Usage Requirements

- Large data set with good quality (input-output mappings)
- Measurable and describable goals (define the cost)
- Enough computing power (AWS-GPU instance)
- Excels in task where the basic unit (pixels, words) has very little meaning in itself, but the combination of such units has a useful meaning.

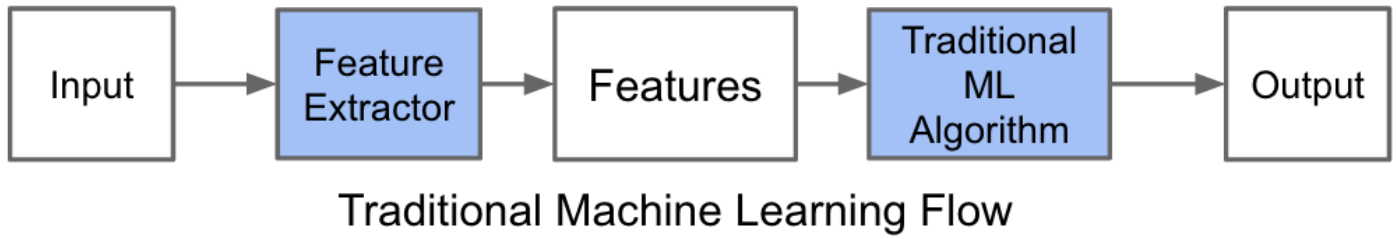
- Large data set with good quality (input-output mappings)
- Measurable and describable goals (define the cost)
- Enough computing power (AWS-GPU instance)
- Excels in task where the basic unit (pixels, words) has very little meaning in itself, but the combination of such units has a useful meaning.

- Large data set with good quality (input-output mappings)
- Measurable and describable goals (define the cost)
- Enough computing power (AWS-GPU instance)
- Excels in task where the basic unit (pixels, words) has very little meaning in itself, but the combination of such units has a useful meaning.

The most fundamental difference between deep learning and traditional machine learning is its performance as the scale of data increases.



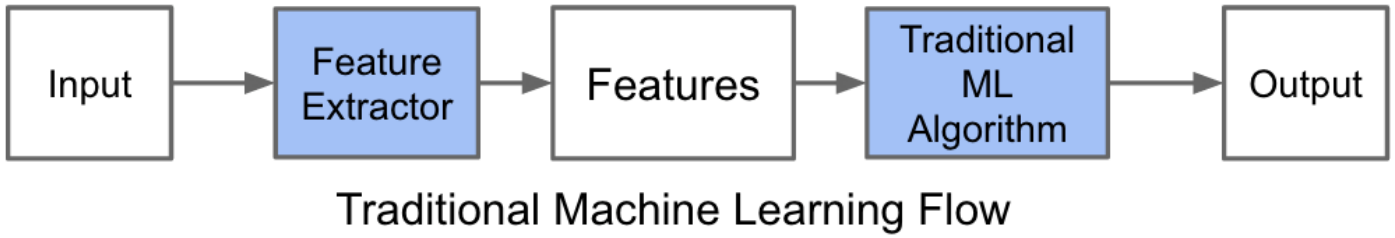
- In Machine learning, most of the applied features need to be identified by an expert and then hand-coded as per the domain and data type.
- Deep learning algorithms try to learn high-level features from data. Therefore, deep learning reduces the task of developing new feature extractor for every problem.



Traditional Machine Learning Flow

Deep Learning Flow

- In Machine learning, most of the applied features need to be identified by an expert and then hand-coded as per the domain and data type.
- Deep learning algorithms try to learn high-level features from data. Therefore, deep learning reduces the task of developing new feature extractor for every problem.



Traditional Machine Learning Flow

Deep Learning Flow

- A deep learning algorithm takes a long time to train. For e.g state of the art deep learning algorithm: ResNet takes about two weeks to train completely from scratch.
- Whereas machine learning comparatively takes much less time to train, ranging from a few seconds to a few hours.

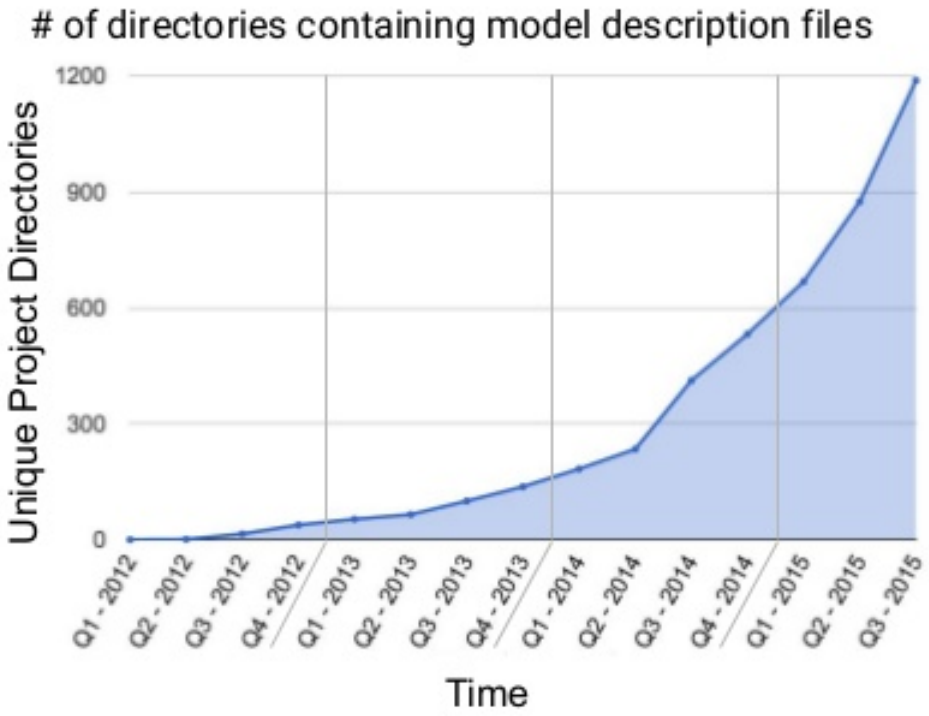
- A deep learning algorithm takes a long time to train. For e.g state of the art deep learning algorithm: ResNet takes about two weeks to train completely from scratch.
- Whereas machine learning comparatively takes much less time to train, ranging from a few seconds to a few hours.

- At test time, deep learning algorithm takes much less time to run.
- Whereas, if you compare machine learning algorithms, test time generally increases on increasing the size of data.

- At test time, deep learning algorithm takes much less time to run.
- Whereas, if you compare machine learning algorithms, test time generally increases on increasing the size of data.

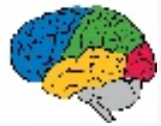

THE POSSIBILITIES ARE ENDLESS.

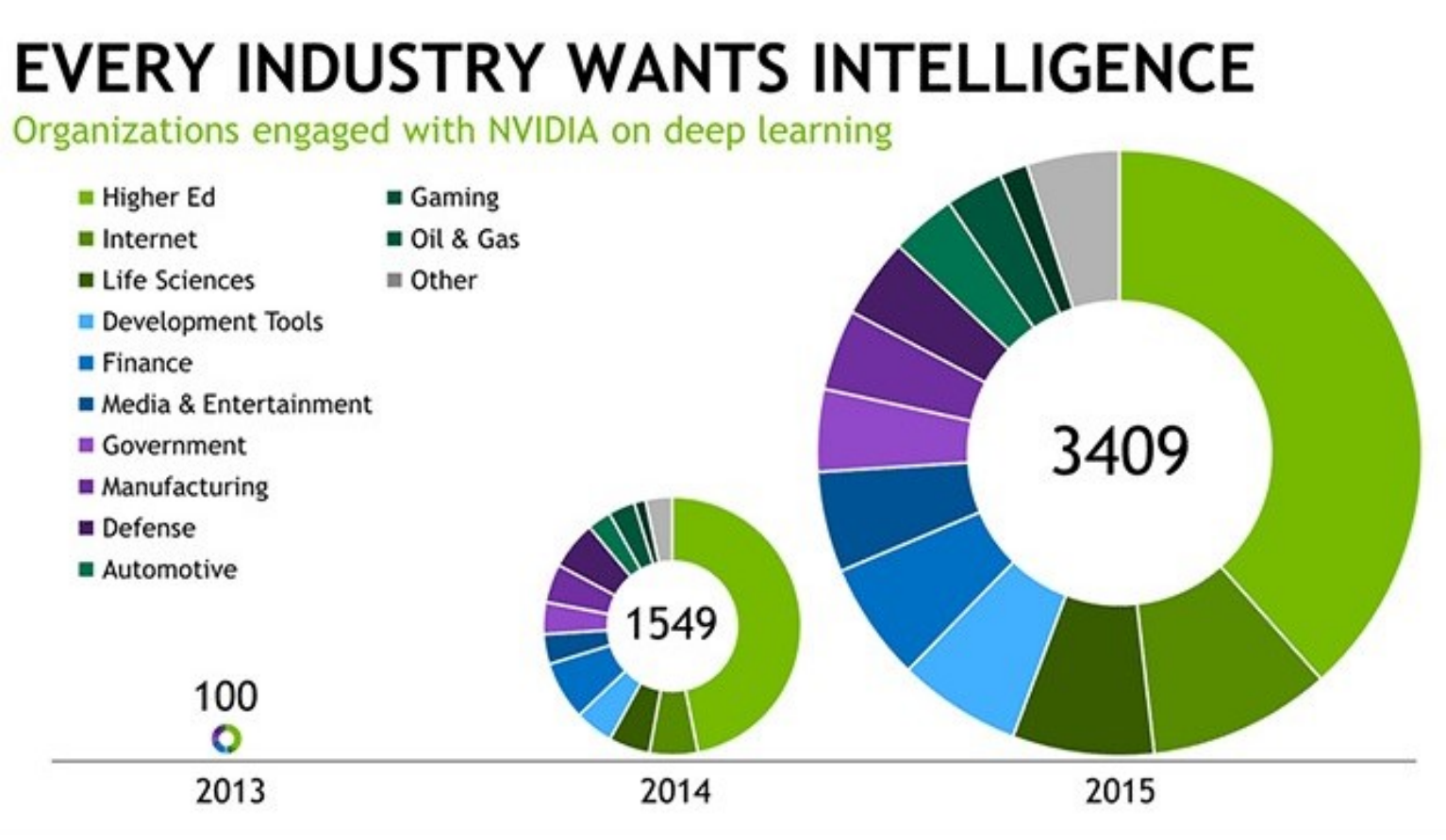
Growing Use of Deep Learning at Google



Across many products/areas:

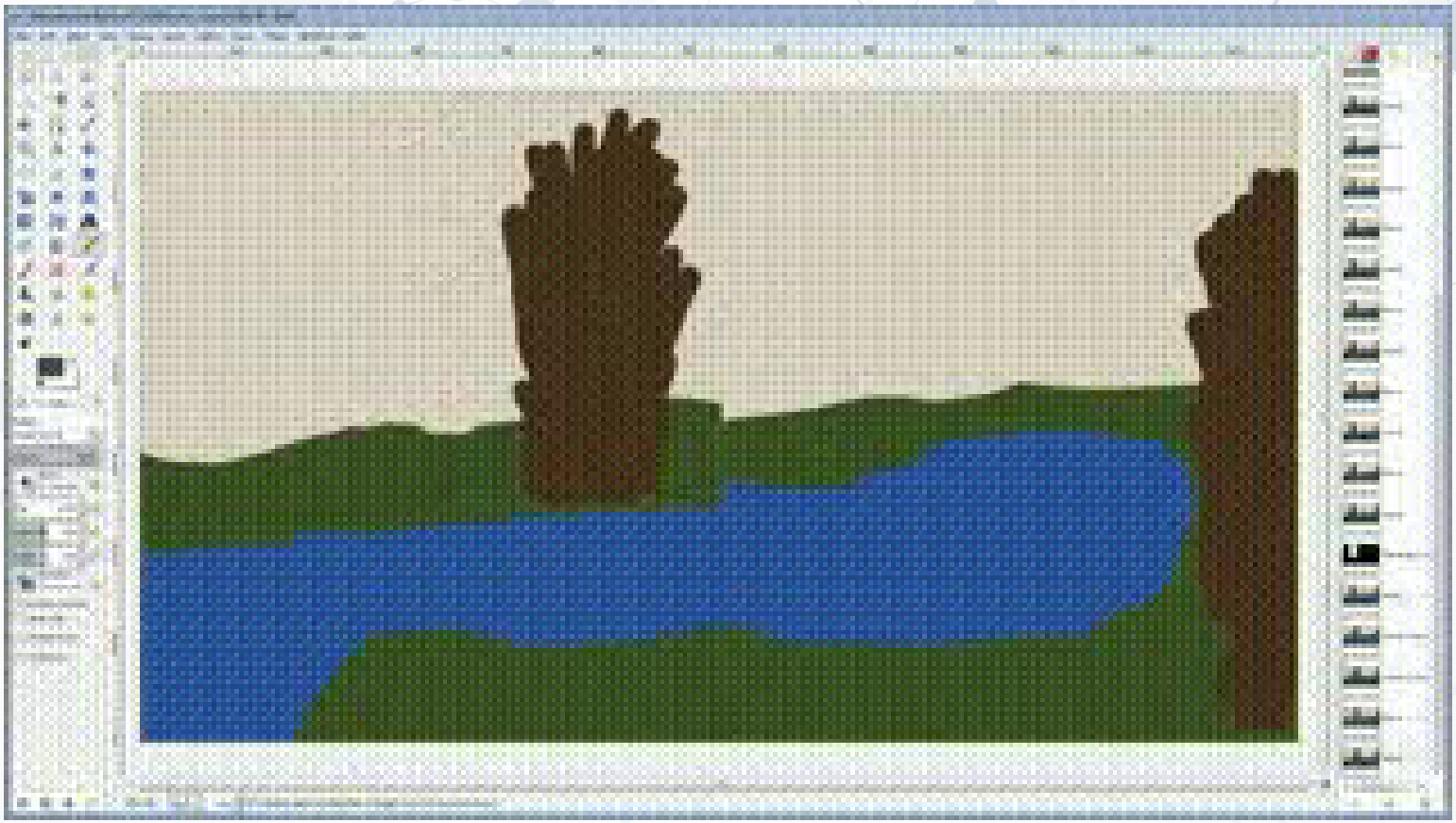
Android
Apps
drug discovery
Gmail
Image understanding
Maps
Natural language understanding
Photos
Robotics research
Speech
Translation
YouTube
… many others …





- **LipNet:**Automatic lip-reading implemented via deep learning, having accuracy far better than humans.

- **Neural Doodle:** Converts 2D doodles into artistic masterpieces.



- Automatic Image Captioning

A person on a beach flying a kite.

A black and white photo of a train on a train track.

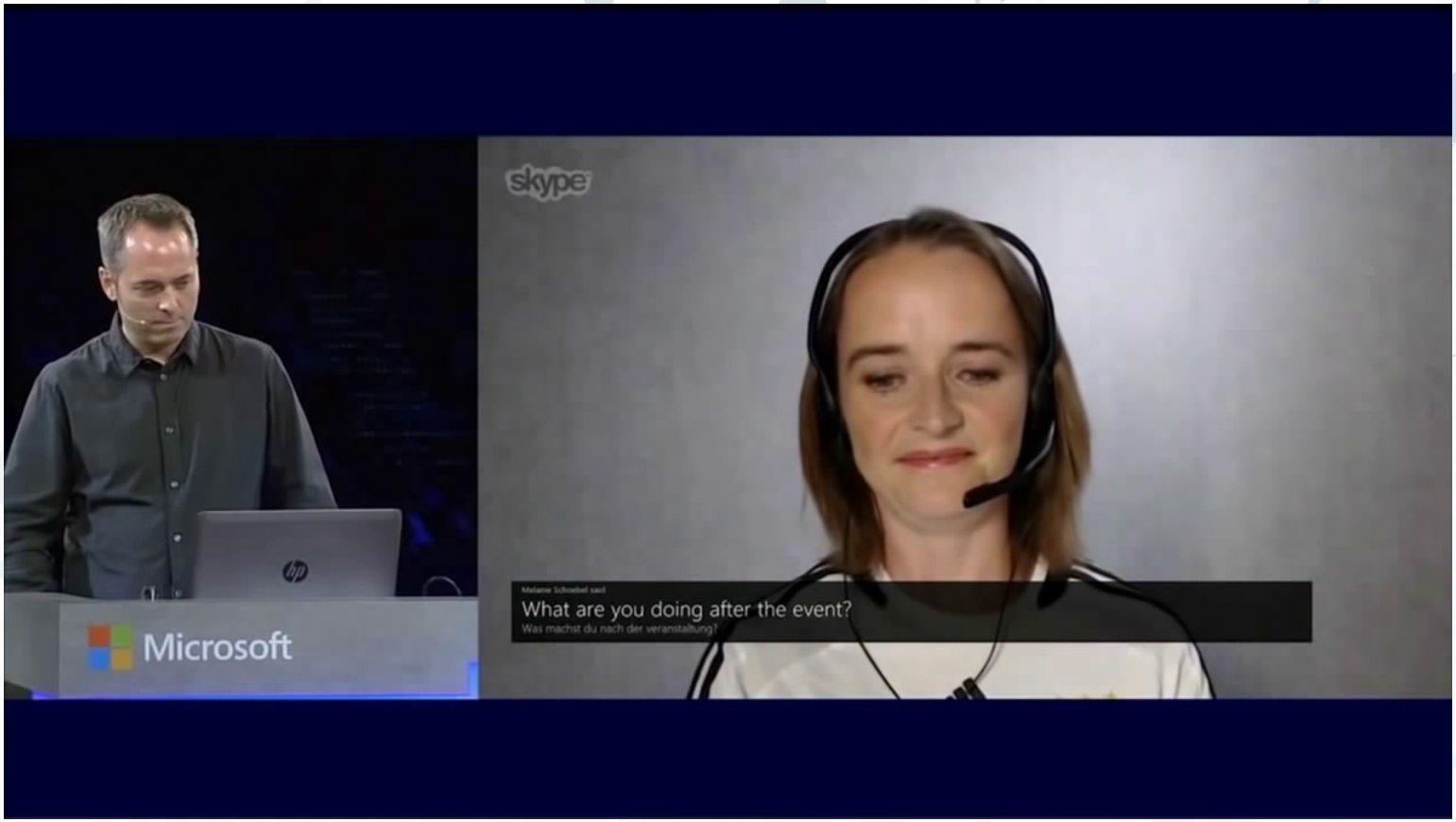
- Automatic Image Captioning

A person skiing down a snow covered slope.

A group of giraffe standing next to each other.

- Automatic Image Captioning

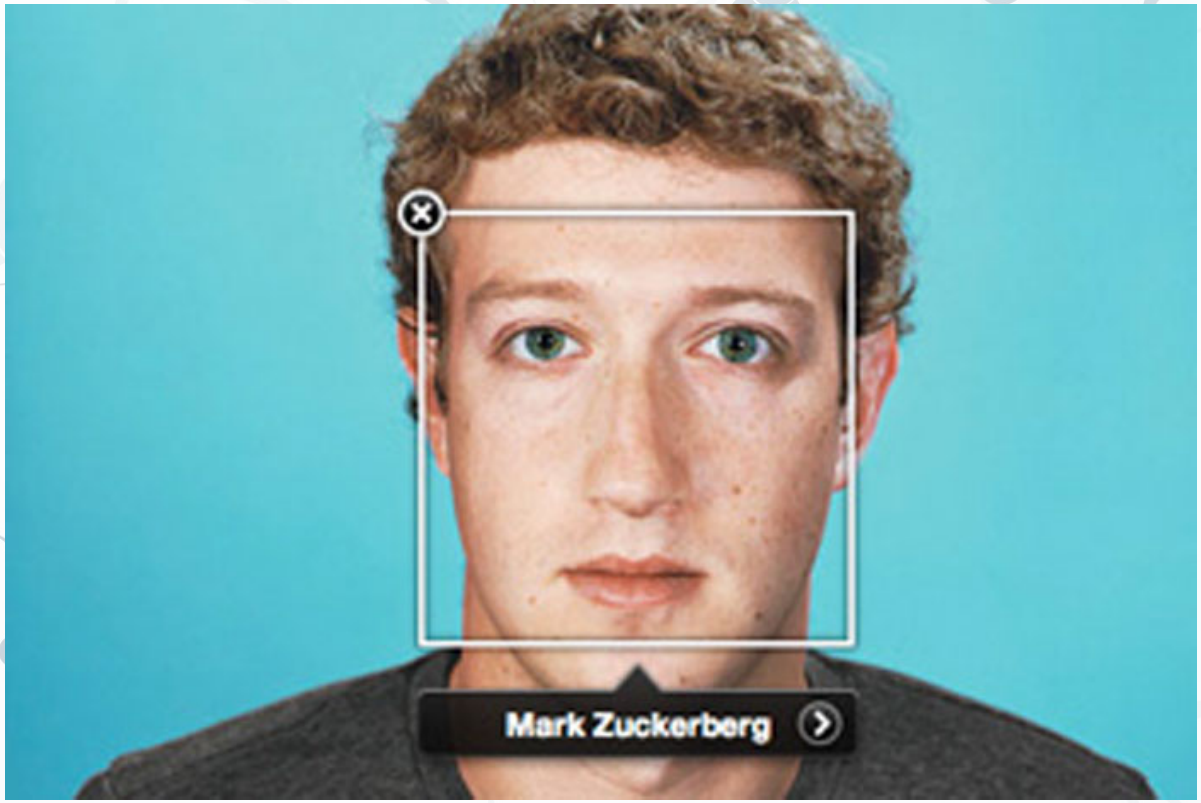
- Skype's real time translation



- Prisma App's artistic rendering of photographs

- Facebook for blind.

- Facebook's automatic facial recognition and photo-tagging.



- Facebook's **DeepFace**

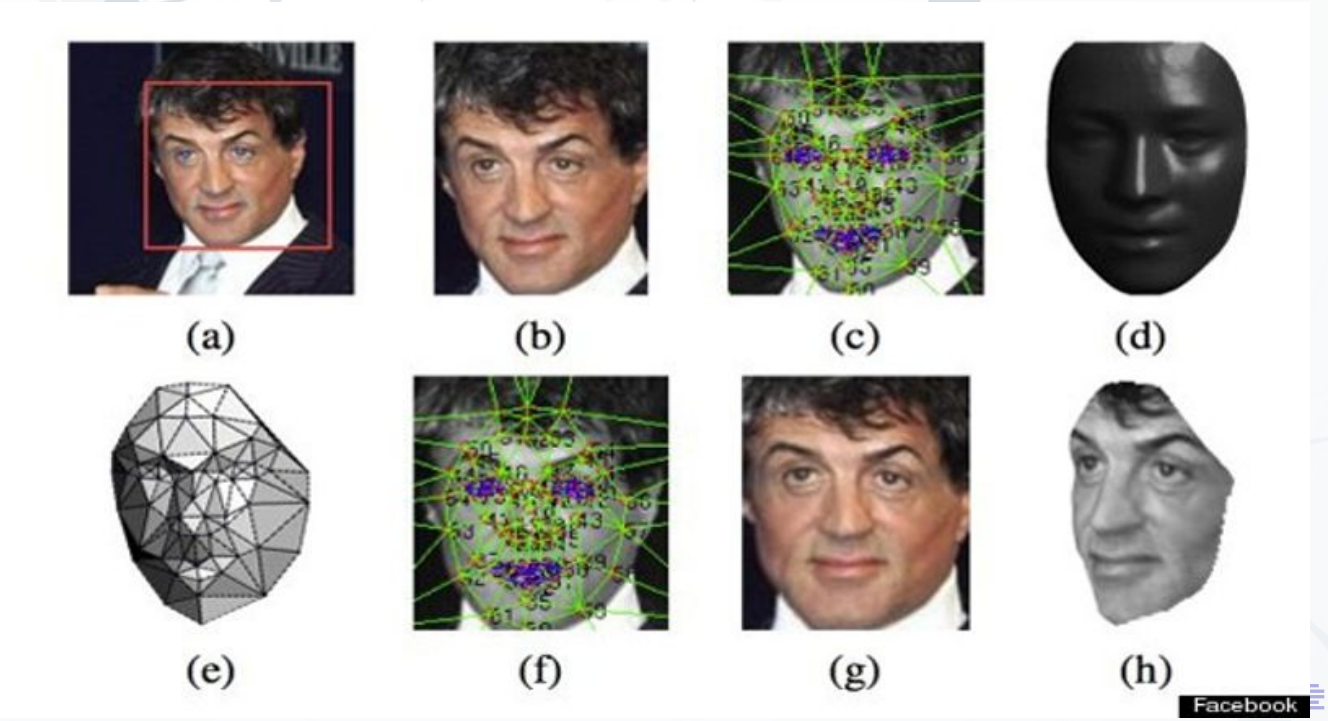
Other face recognition models perform these steps:- detect - align - represent - classify

Deep Face employs 3-dimensional modeling to revisit both the alignment and representation step.

- Facebook's **DeepFace**

Other face recognition models perform these steps:- detect - align - represent - classify

Deep Face employs 3-dimensional modeling to revisit both the alignment and representation step.



- Facebook's **DeepFace**

Other face recognition models perform these steps:- detect - align - represent - classify

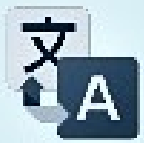
Deep Face employs 3-dimensional modeling to revisit both the alignment and representation step.

It employs a nine-layer neural net with over 120 million connection weights, and was trained on four million images uploaded by Facebook users. The system is said to be 97 per cent accurate that is quite ahead as compared to 85 per cent of the FBI's.

- Deep Learning in Google products

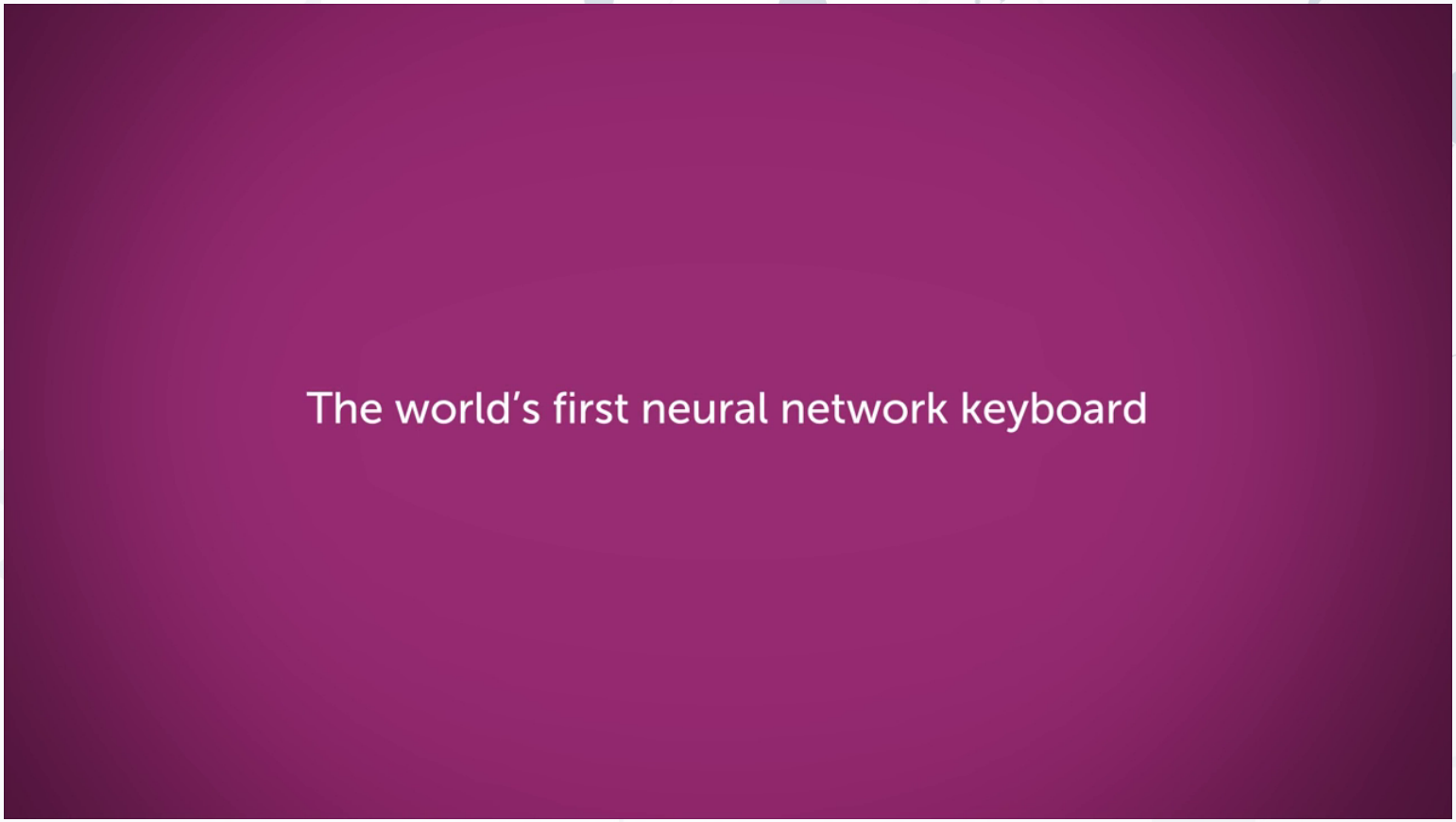
RankBrain (Search): In few months, RankBrain has become the third-most important signal contributing to the search result.

Speech Recognition (Google Now): 30% reduction in Word Error Rate for English. Biggest single improvement in 20 years of speech research.

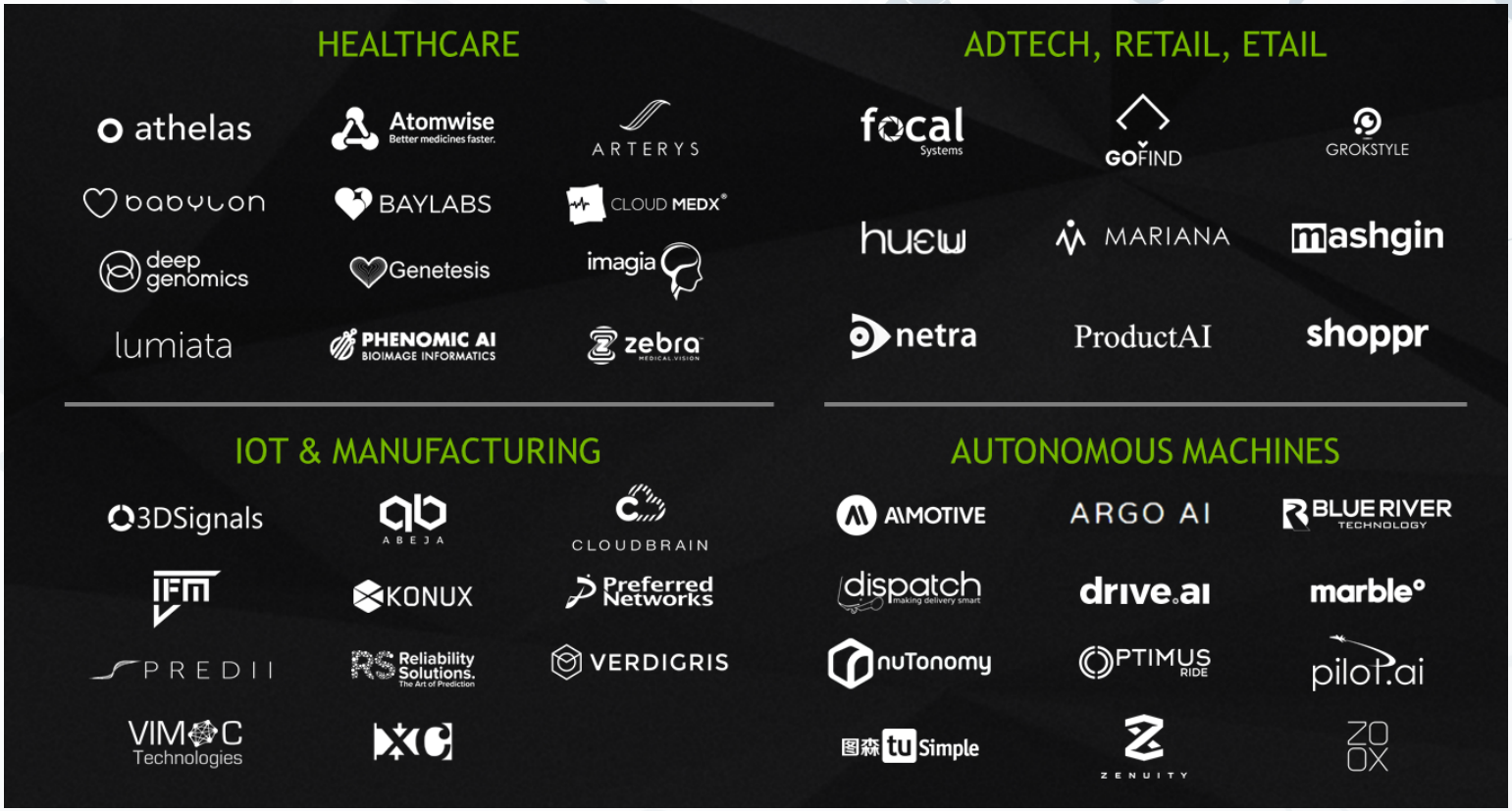


Visual Translation (Google Translate): Real-time visual translation of 20 more languages all on the phone and without an Internet connection using a deep neural net.

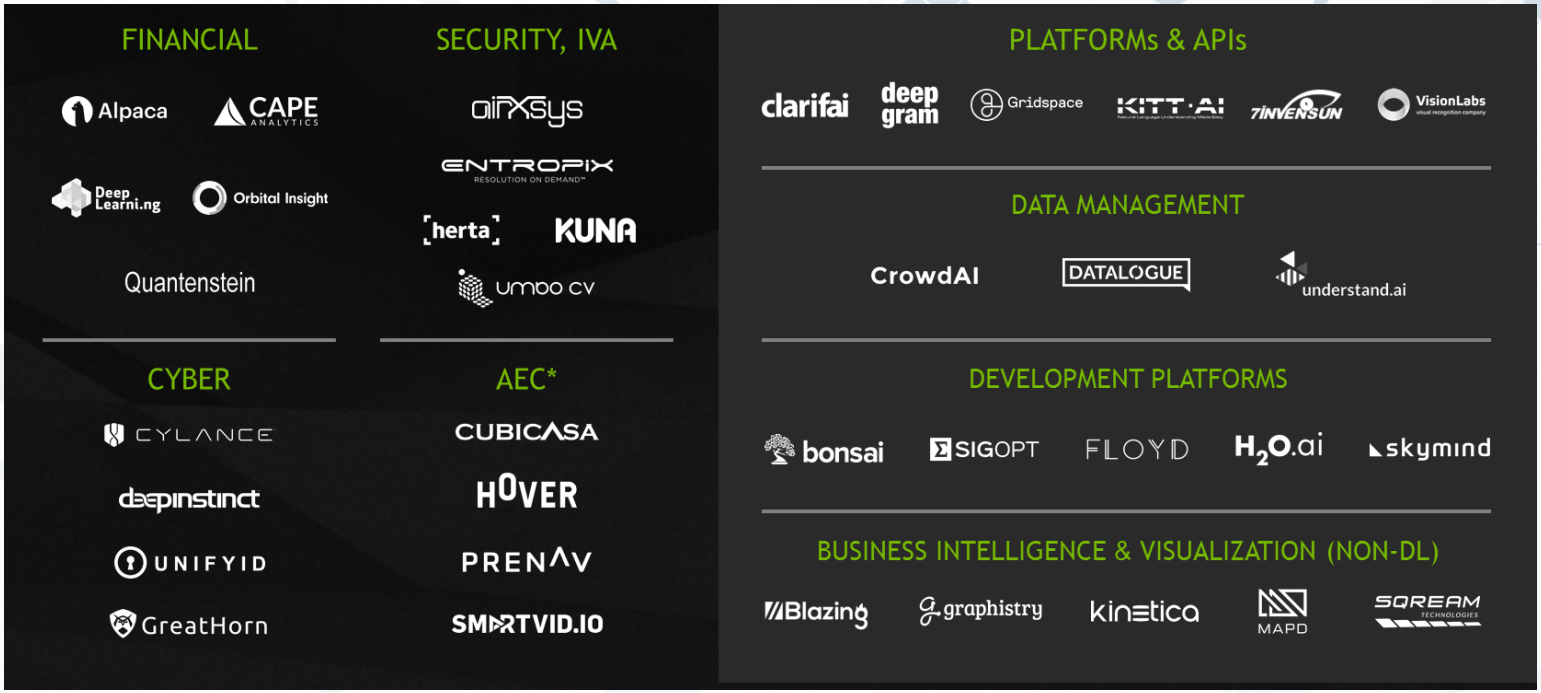
- Swiftkey's Neural Alpha Keyboard



- Over 1300 startups under the **NVIDIA Inception** program



- Over 1300 startups under the **NVIDIA Inception** program



Deeper into Deep Learning

Neural Network: How similar is it to the human brain?

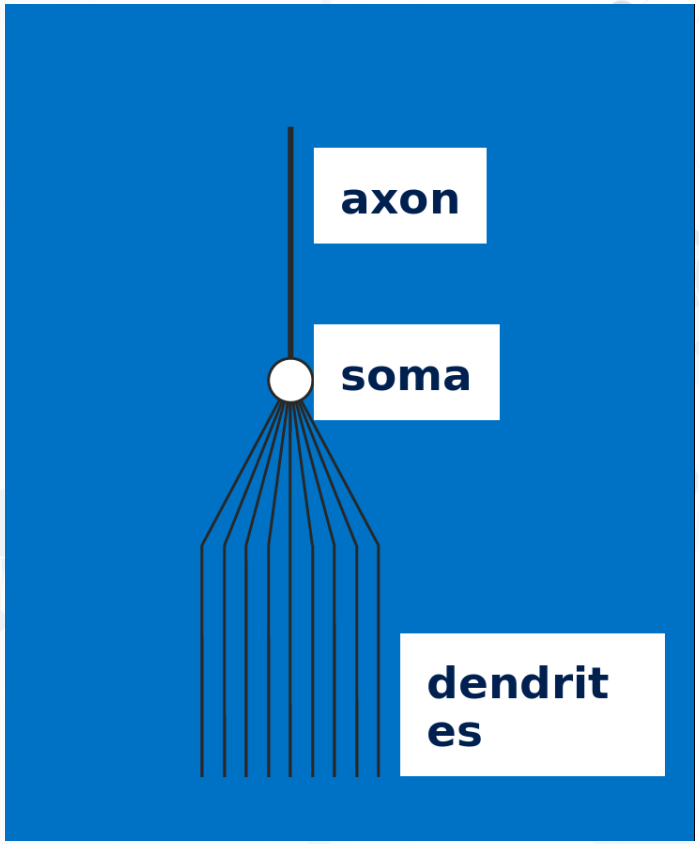


Figure: ANN

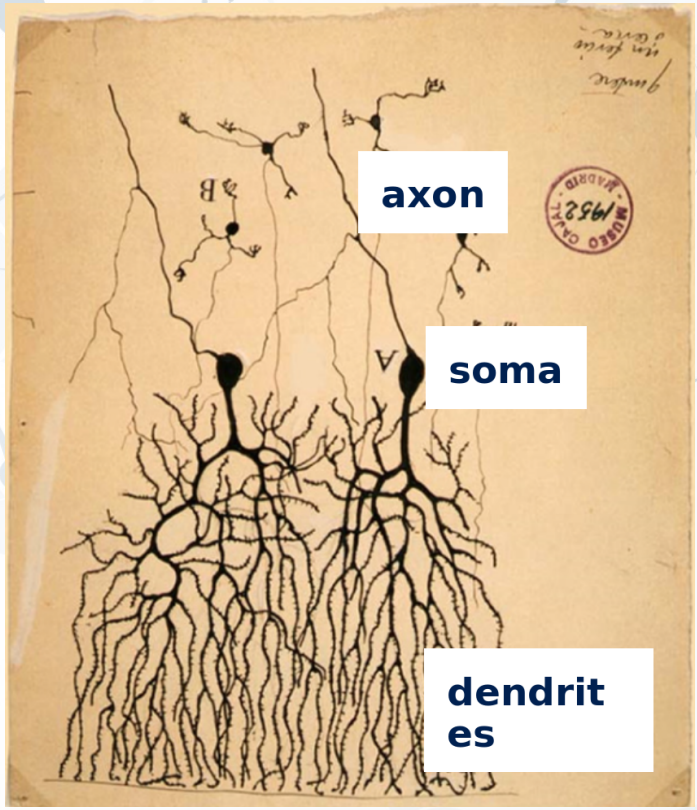
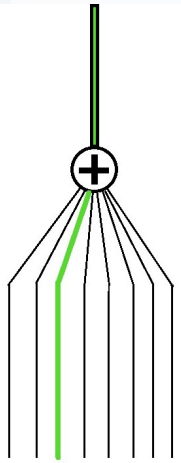
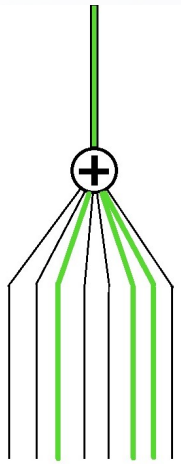


Figure: Human Neural Network

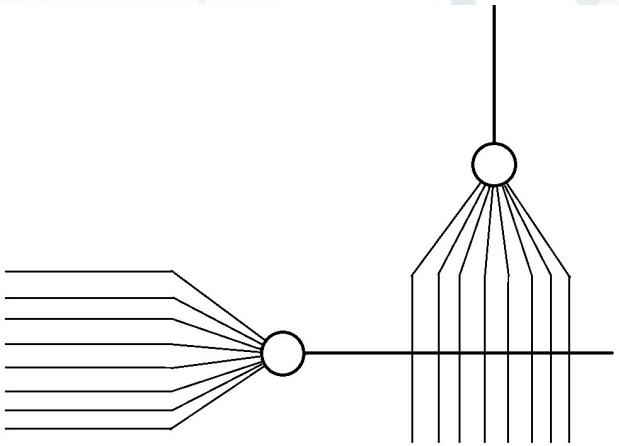
Soma adds dendrite activity together and passes it to axon.

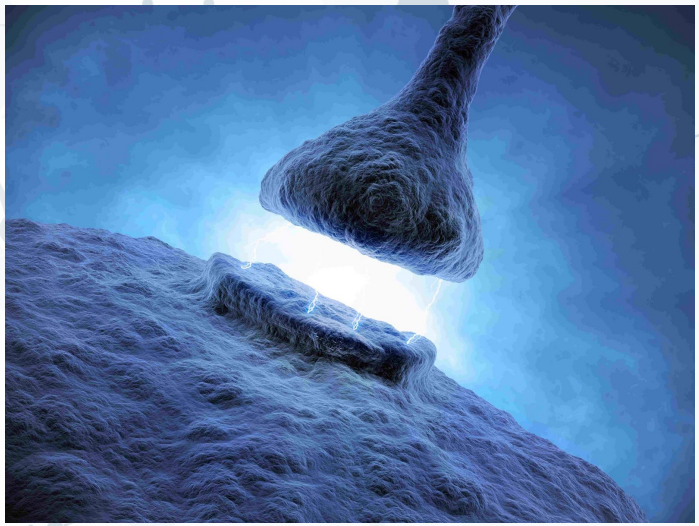


More dendrite activity makes more axon activity.

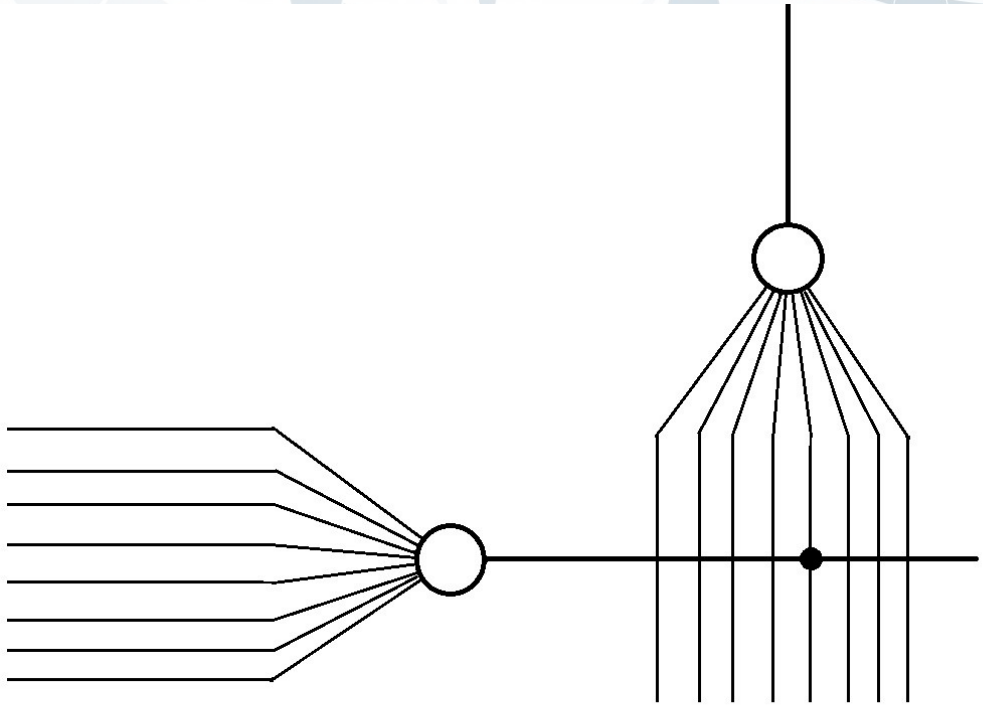


Synapse: ”connection between axon of one neurons and dendrites of another”

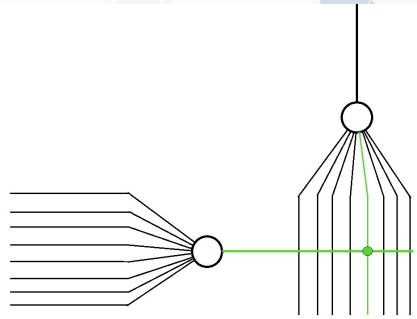




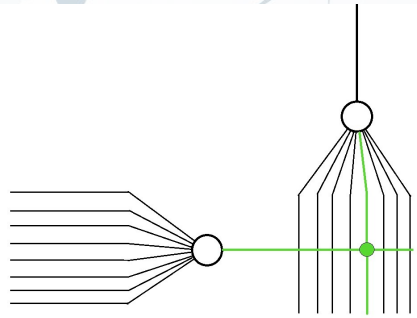
Axons can connect to dendrites strongly, weakly, or somewhere in between.



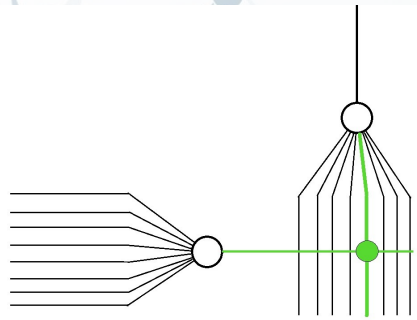
Weak Connection (0.2)



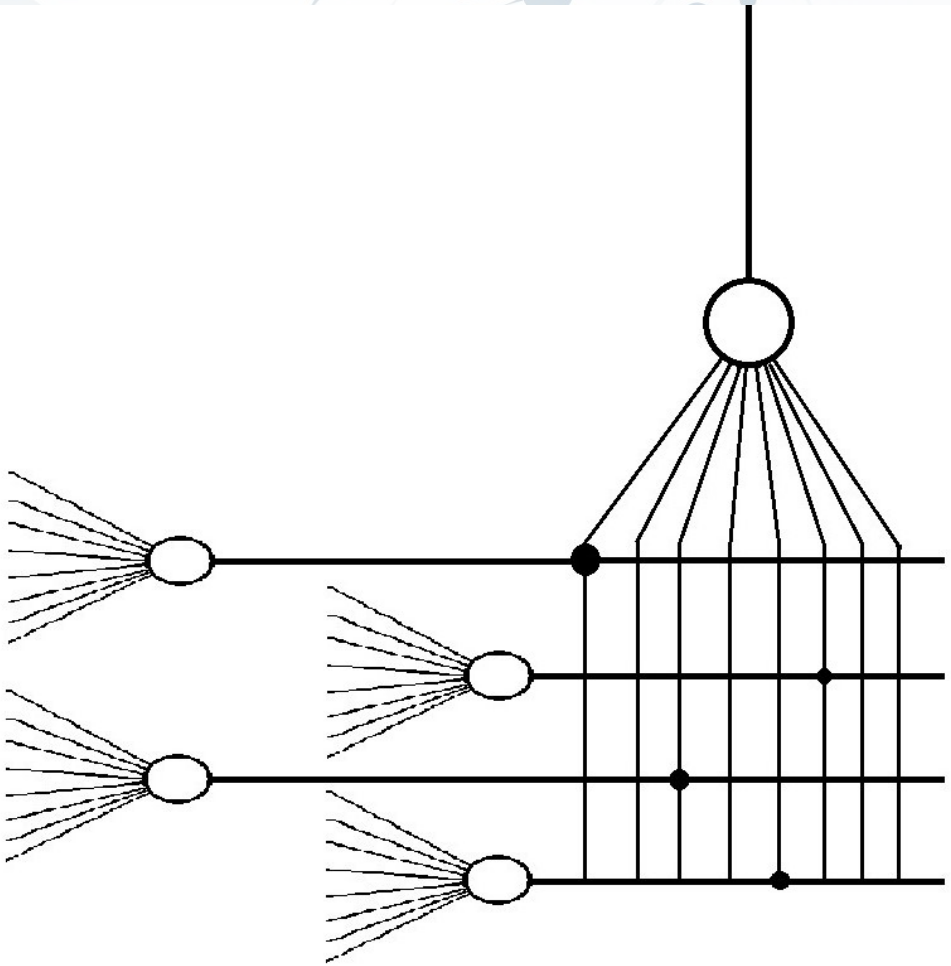
Medium Connection (0.6)



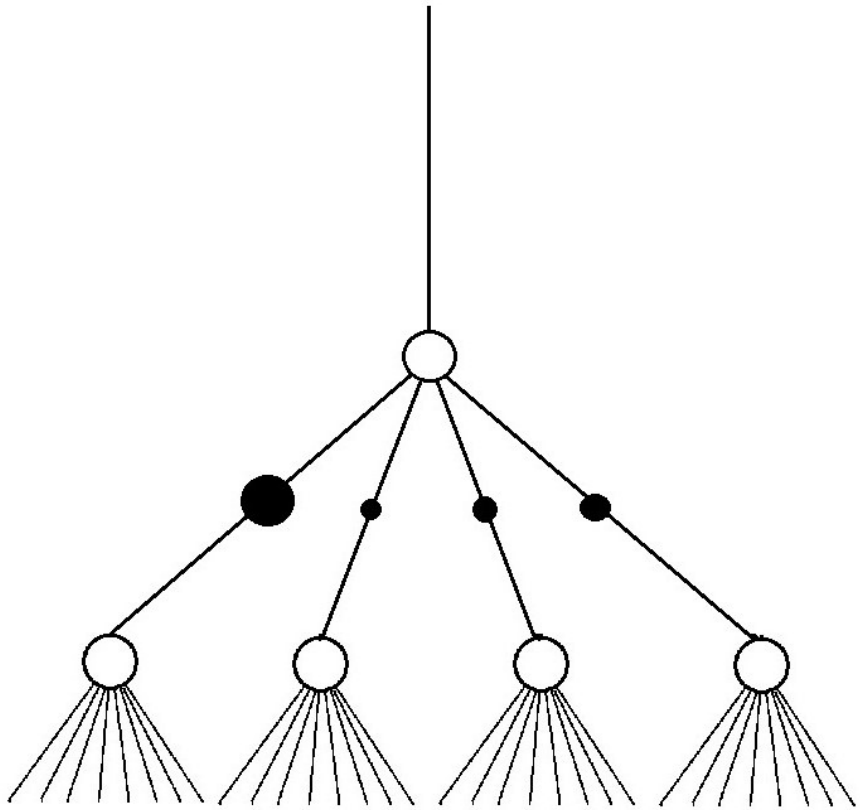
Strong Connection (1.0)



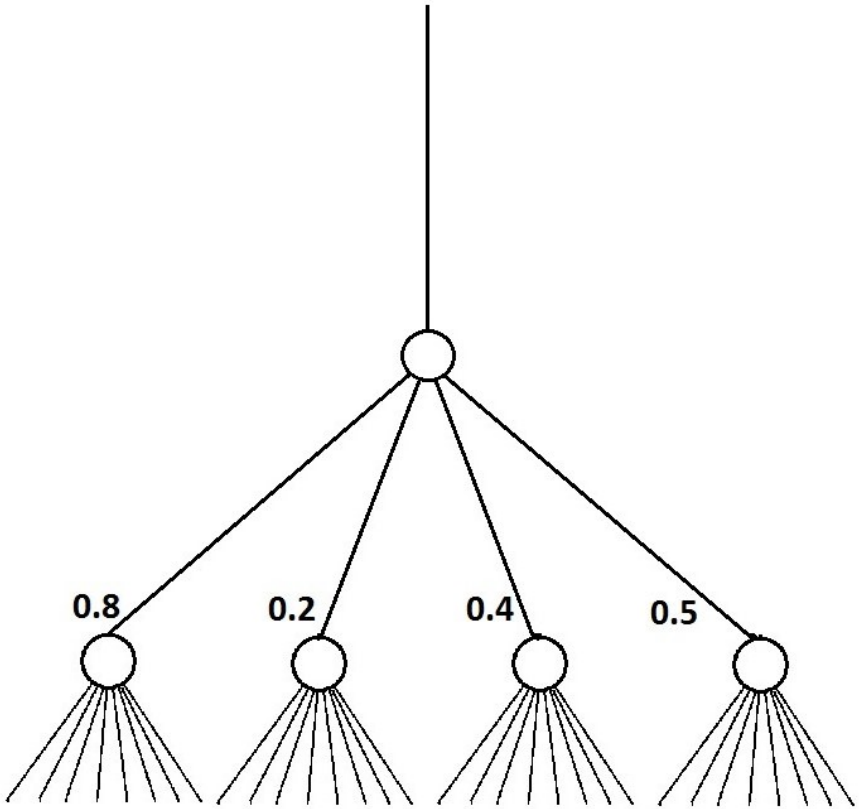
Lots of axons connect with dendrites of one neuron.Each has its own connection strength.



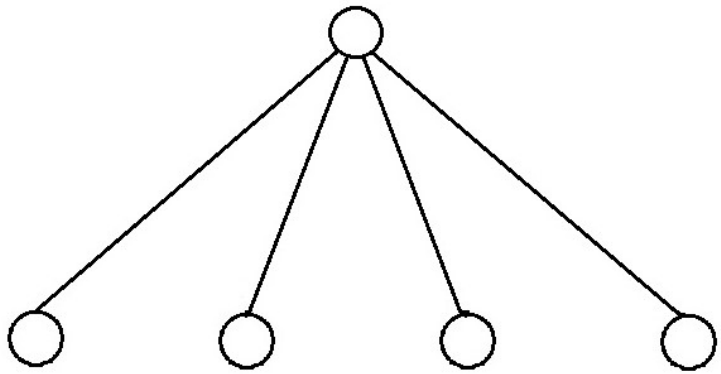
The above illustration can be simplified as above.



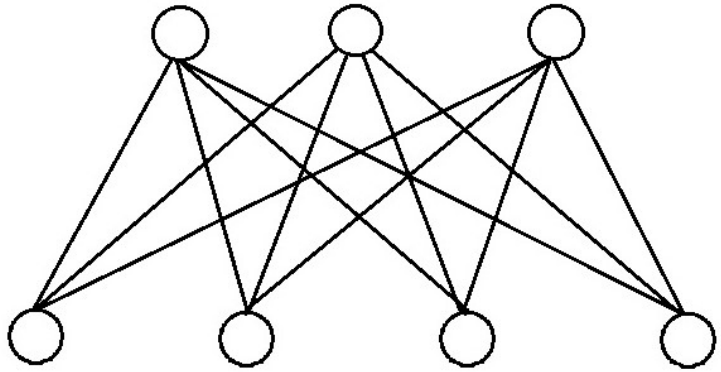
On giving numerical values to the strength of connections i.e. weights



A much simplified version looks something like this.

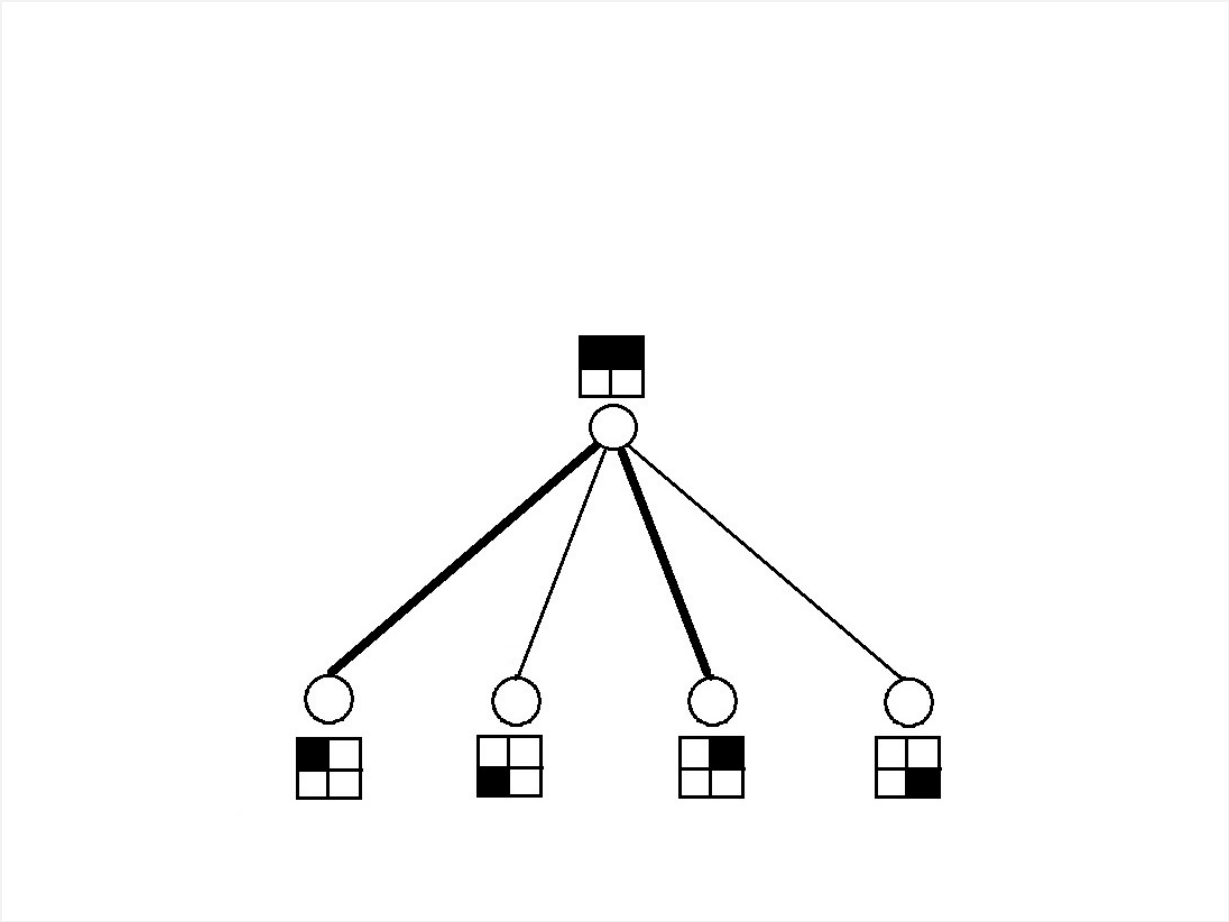


On increasing the number of neurons and synapses.

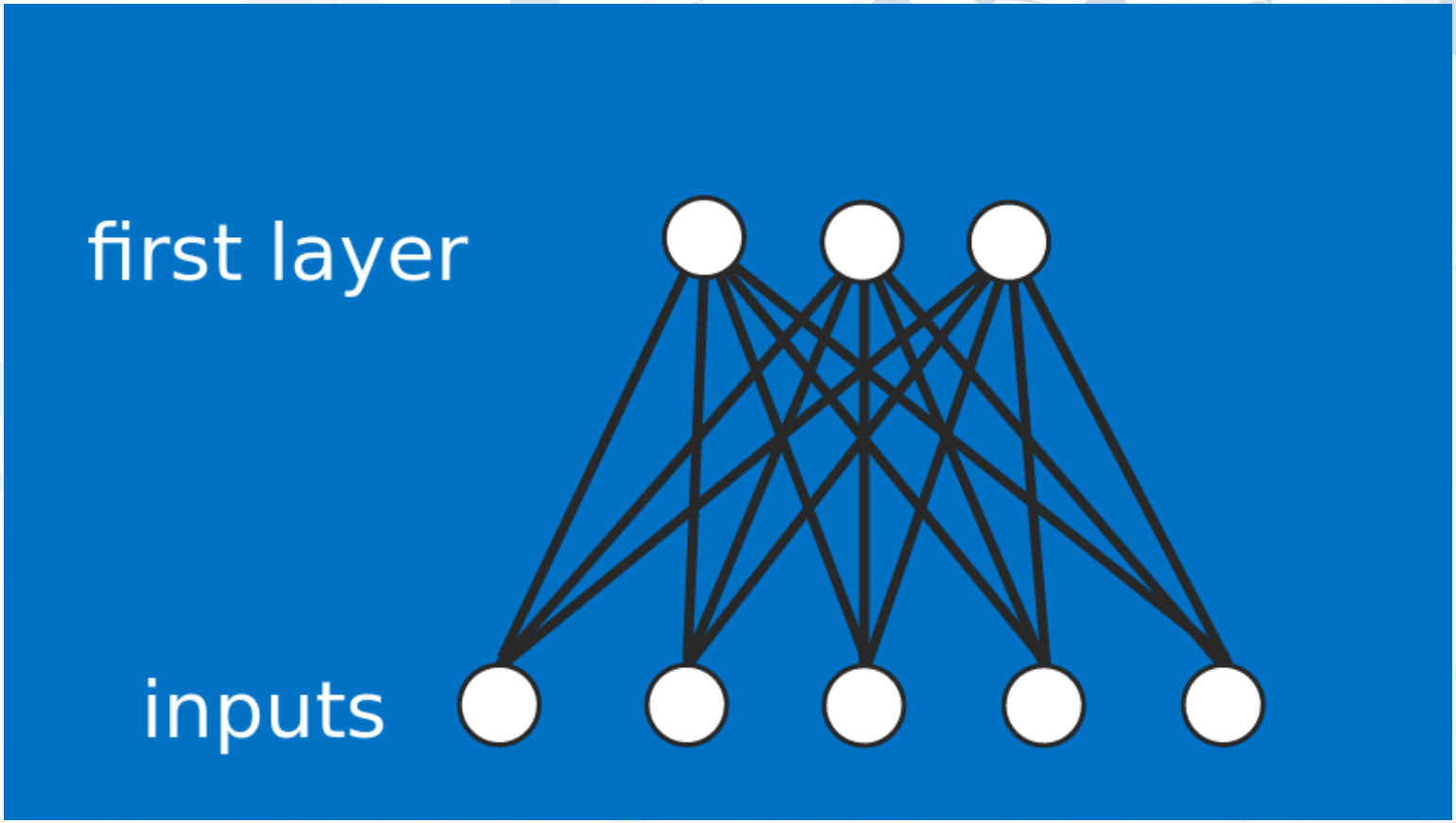


An example

Suppose the first and third input has been activated.



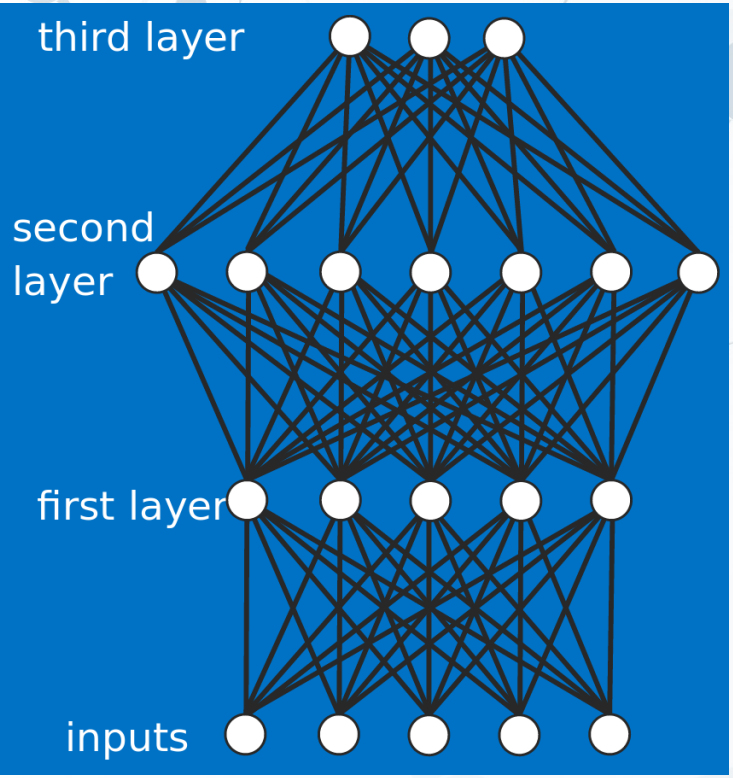
Each node represents a pattern, a combination of neurons of the previous layers.



Deep network

If a network has more than three layers, it's deep.

Some 12 or more.



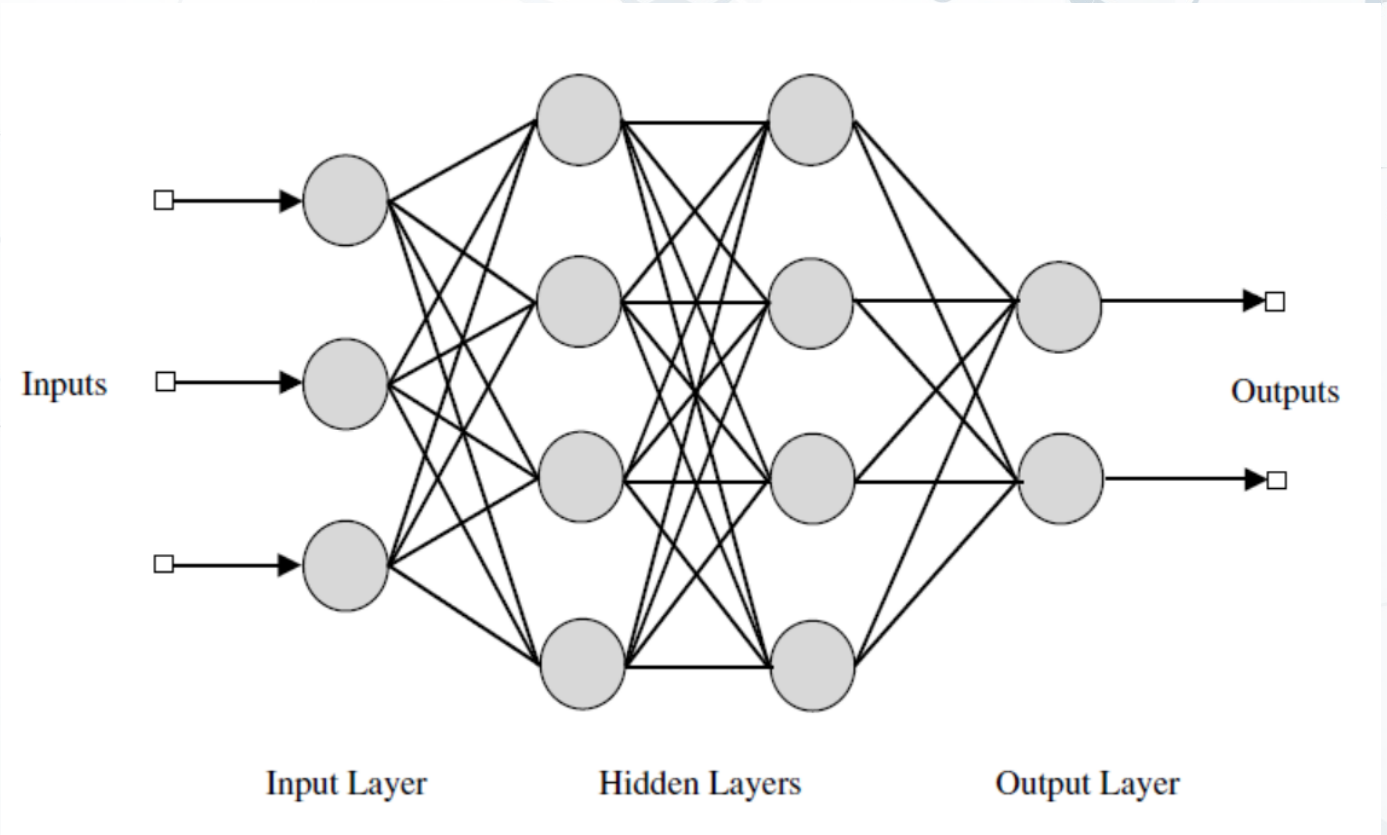
Two Main types of artificial neural networks are:-

- Feedforward Neural Network
- Recurrent Neural Network (RNN)

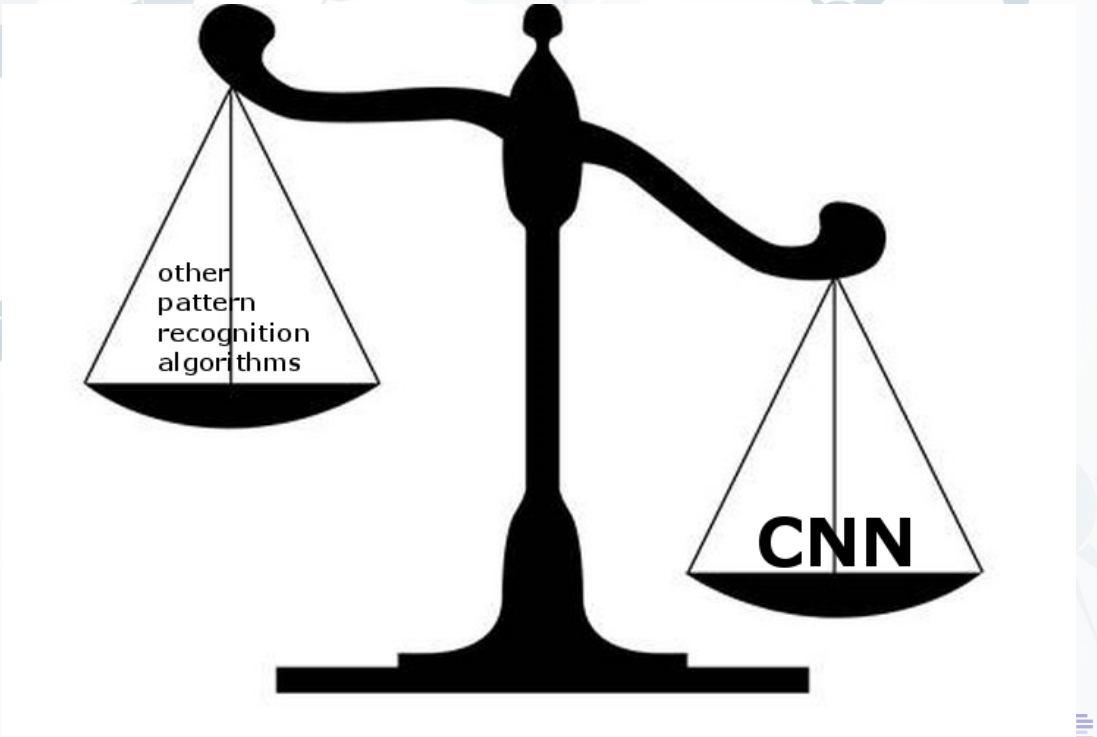
- **Feedforward Neural Network**
  - Convolutional neural network (CNN)
  - Autoencoder
  - Probabilistic neural network (PNN)
  - Time delay neural network (TDNN)

- **Recurrent Neural Network (RNN)**
  - Long short-term memory RNN (LSTM)
  - Fully recurrent Network
  - Simple recurrent Network
  - Echo state network
  - Bi-directional RNN
  - Hierarchical RNN
  - Stochastic neural network

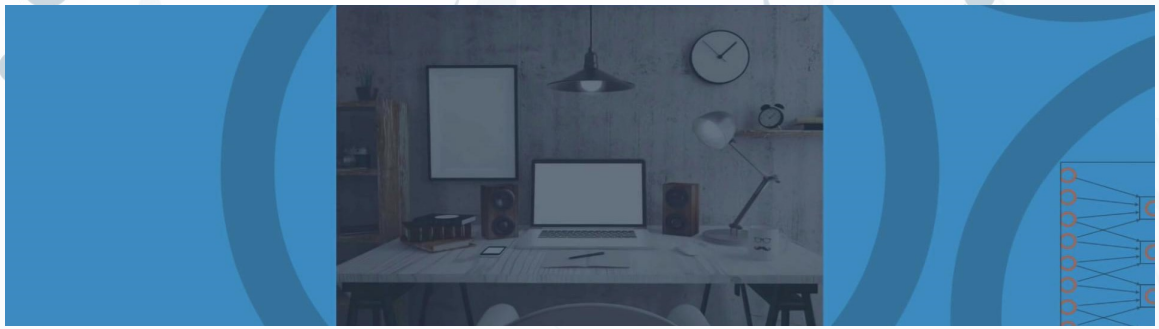
The feedforward neural network was the first and simplest type. In this network the information moves only from the input layer directly through any hidden layers to the output layer without cycles/loops.



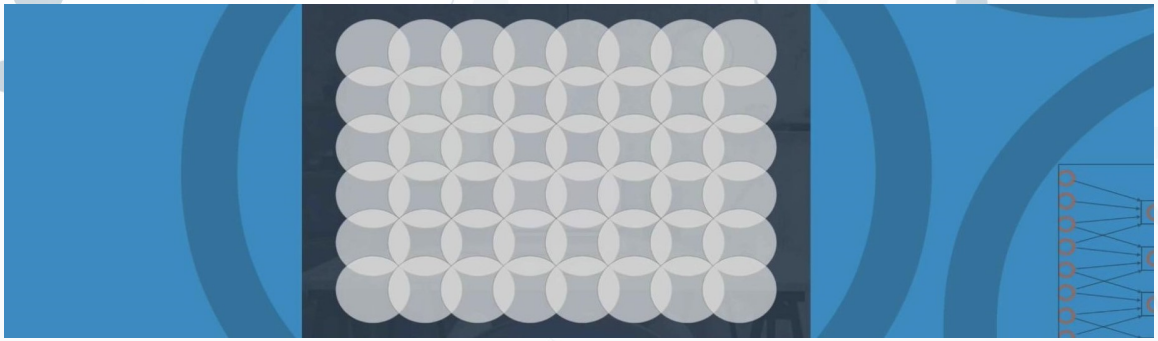
Convolutional Neural Networks learn a complex representation of visual data using vast amounts of data. They are inspired by the human visual system and learn multiple layers of transformations, which are applied on top of each other to extract a progressively more sophisticated representation of the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



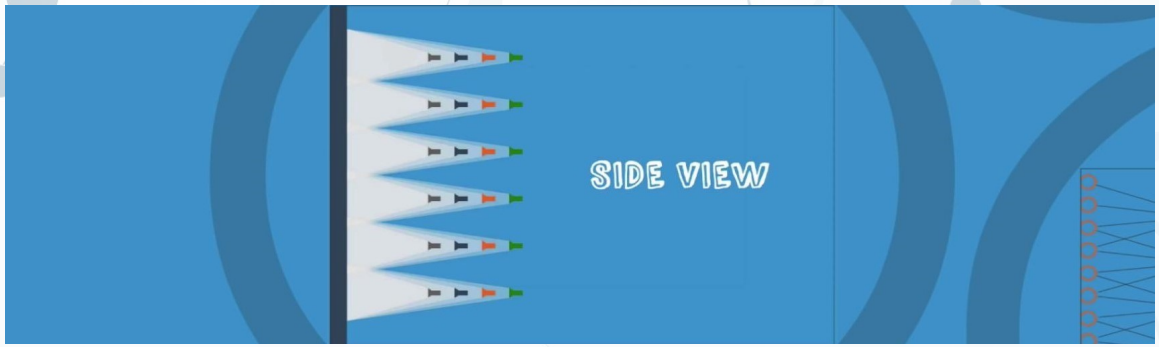
- **Convolutional Layer:** Convolutional Layer is a feature detector that automatically learns to filter out unnecessary information from an input by using convolution kernel.
- **ReLU Layers:** ReLU is the abbreviation ofRectified Linear Units. This layer applies the non-saturatingactivation function $(f(x) = max(0, x))$. It increases the nonlinear properties of the decision function and of the overall network without affecting the receptive fields of the convolution layer.
- **Pooling Layers:** It computes the max or average value of a particular feature over a region of the input data. Also helps to detect objects in some unusual places and reduces memory size.



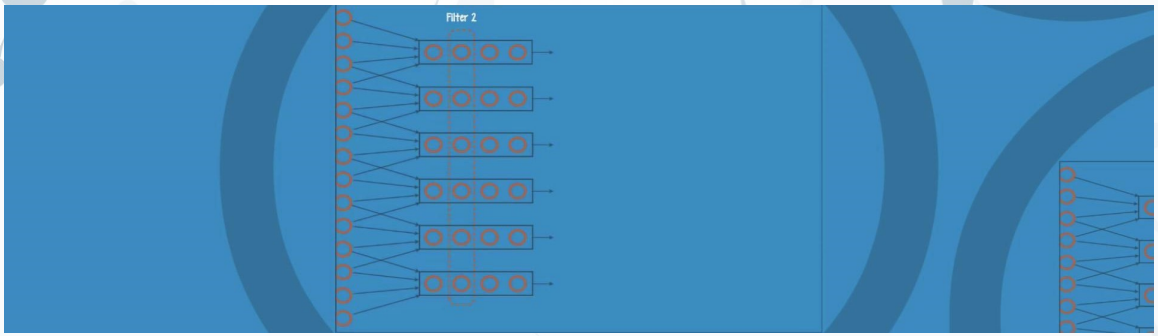
- **Convolutional Layer:** Convolutional Layer is a feature detector that automatically learns to filter out unnecessary information from an input by using convolution kernel.
- **ReLU Layers:** ReLU is the abbreviation ofRectified Linear Units. This layer applies the non-saturatingactivation function $(f(x) = max(0, x))$. It increases the nonlinear properties of the decision function and of the overall network without affecting the receptive fields of the convolution layer.
- **Pooling Layers:** It computes the max or average value of a particular feature over a region of the input data. Also helps to detect objects in some unusual places and reduces memory size.

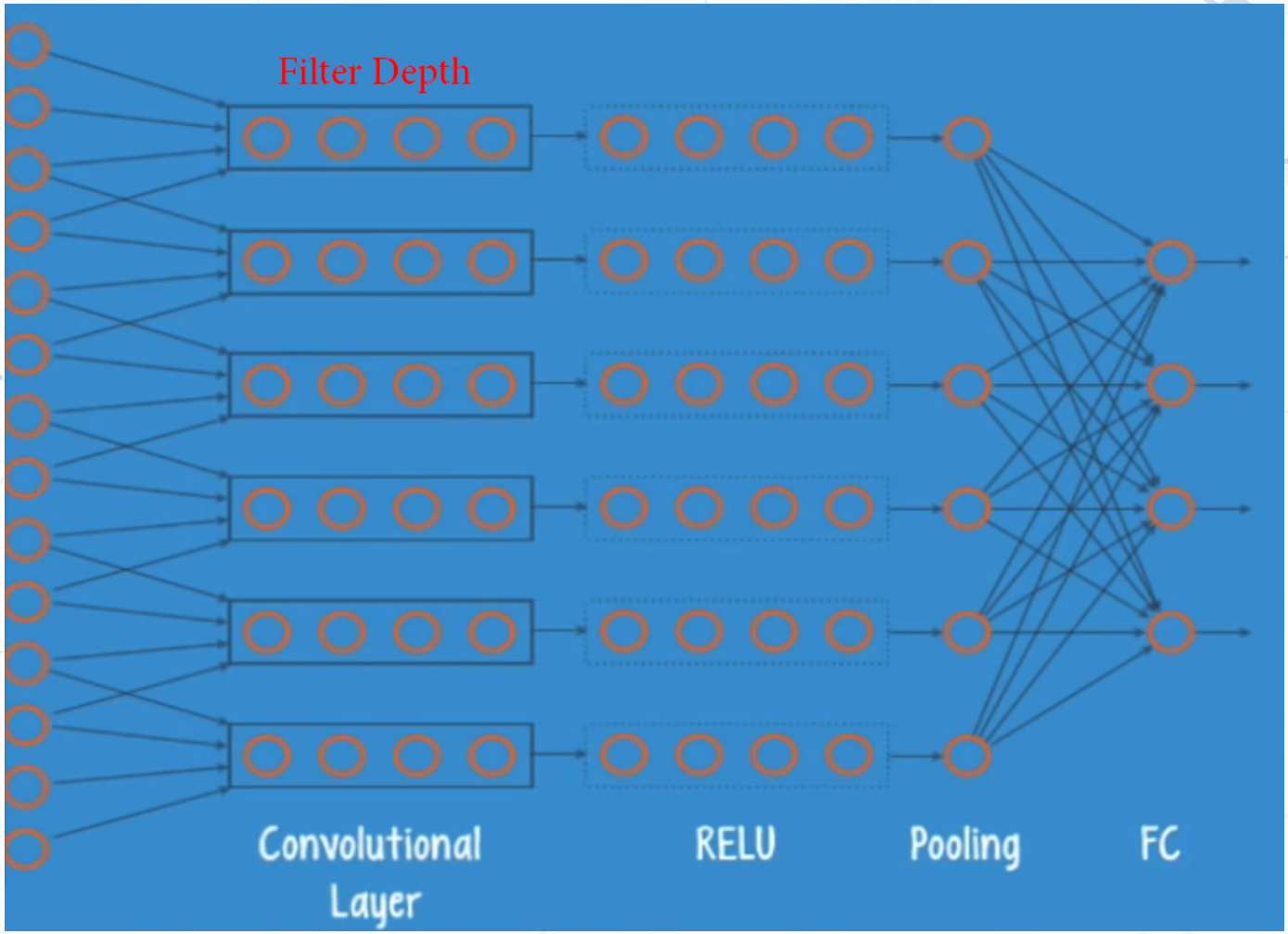


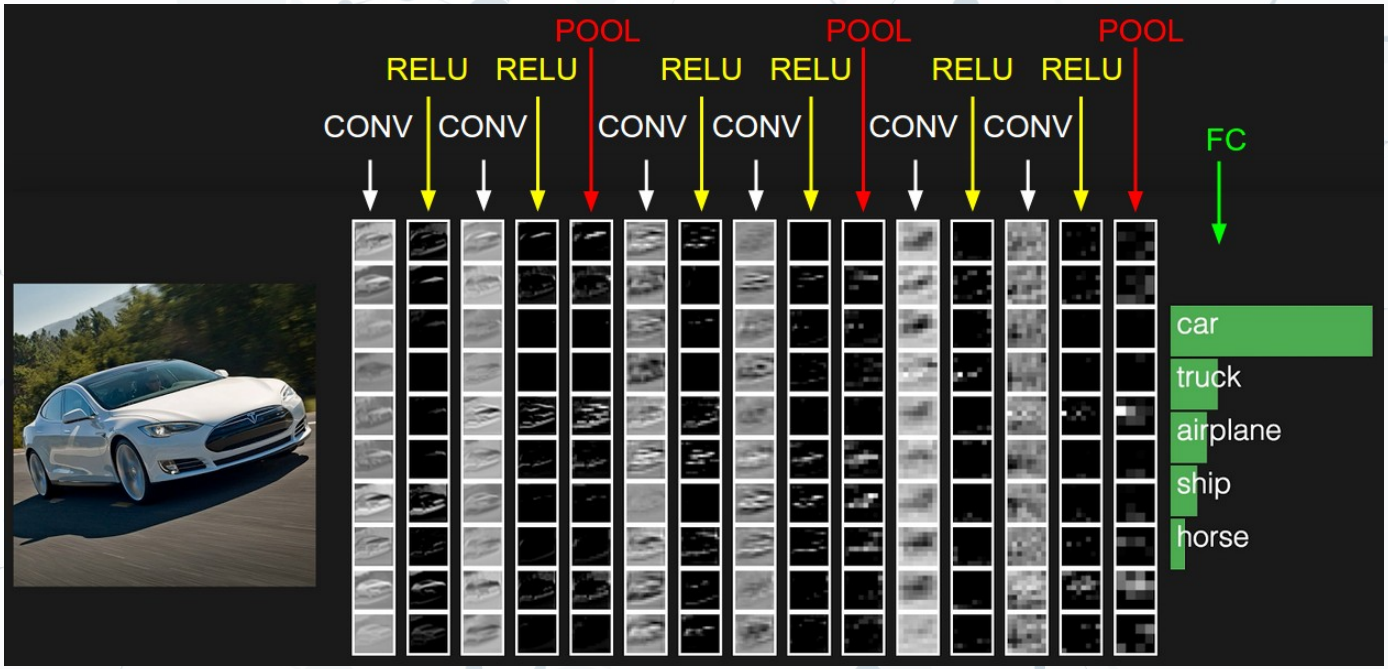
- **Convolutional Layer:** Convolutional Layer is a feature detector that automatically learns to filter out unnecessary information from an input by using convolution kernel.
- **ReLU Layers:** ReLU is the abbreviation ofRectified Linear Units. This layer applies the non-saturatingactivation function $(f(x) = max(0, x))$. It increases the nonlinear properties of the decision function and of the overall network without affecting the receptive fields of the convolution layer.
- **Pooling Layers:** It computes the max or average value of a particular feature over a region of the input data. Also helps to detect objects in some unusual places and reduces memory size.



- **Convolutional Layer:** Convolutional Layer is a feature detector that automatically learns to filter out unnecessary information from an input by using convolution kernel.
- **ReLU Layers:** ReLU is the abbreviation ofRectified Linear Units. This layer applies the non-saturatingactivation function $(f(x) = max(0, x))$. It increases the nonlinear properties of the decision function and of the overall network without affecting the receptive fields of the convolution layer.
- **Pooling Layers:** It computes the max or average value of a particular feature over a region of the input data. Also helps to detect objects in some unusual places and reduces memory size.

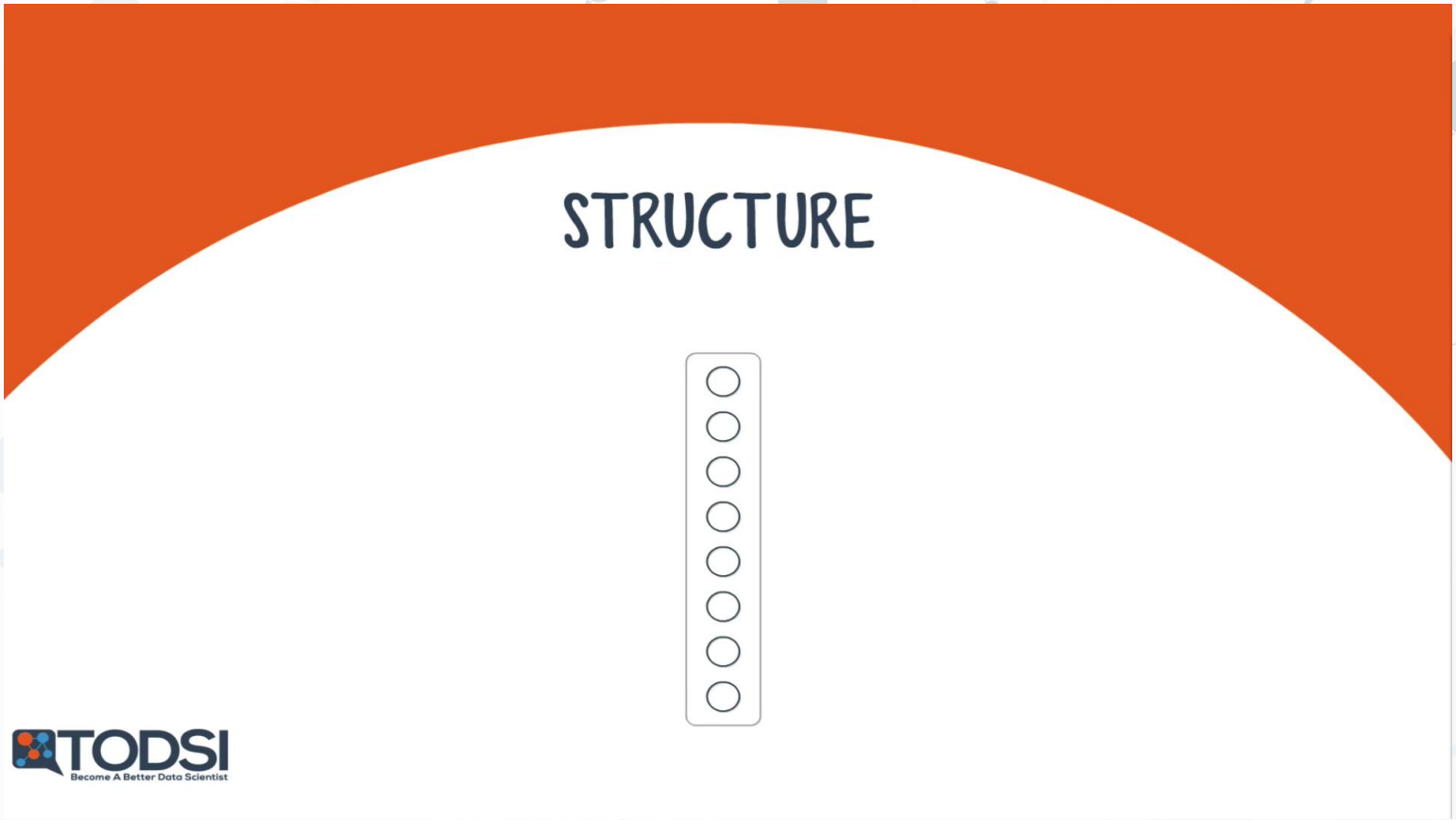




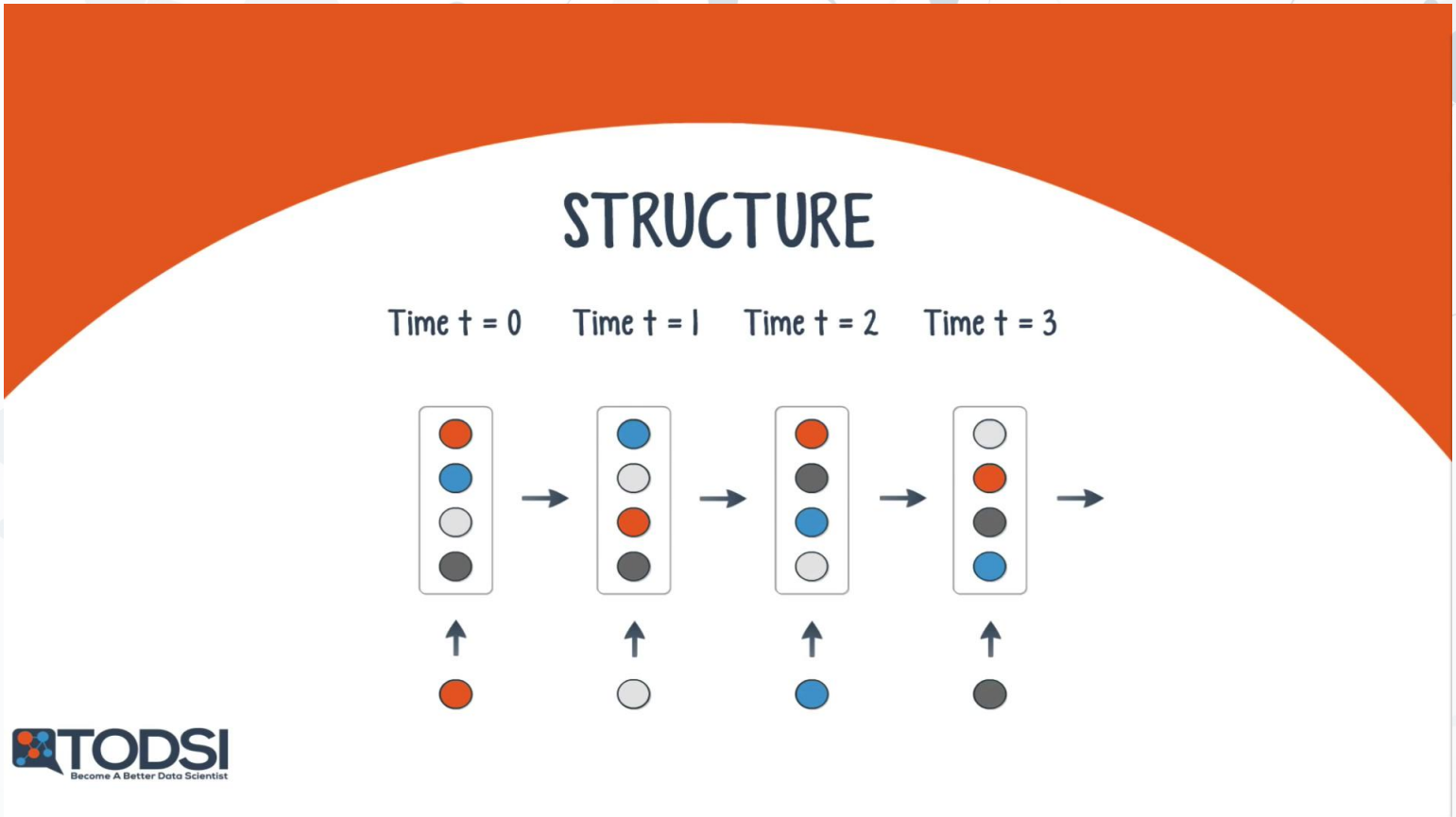


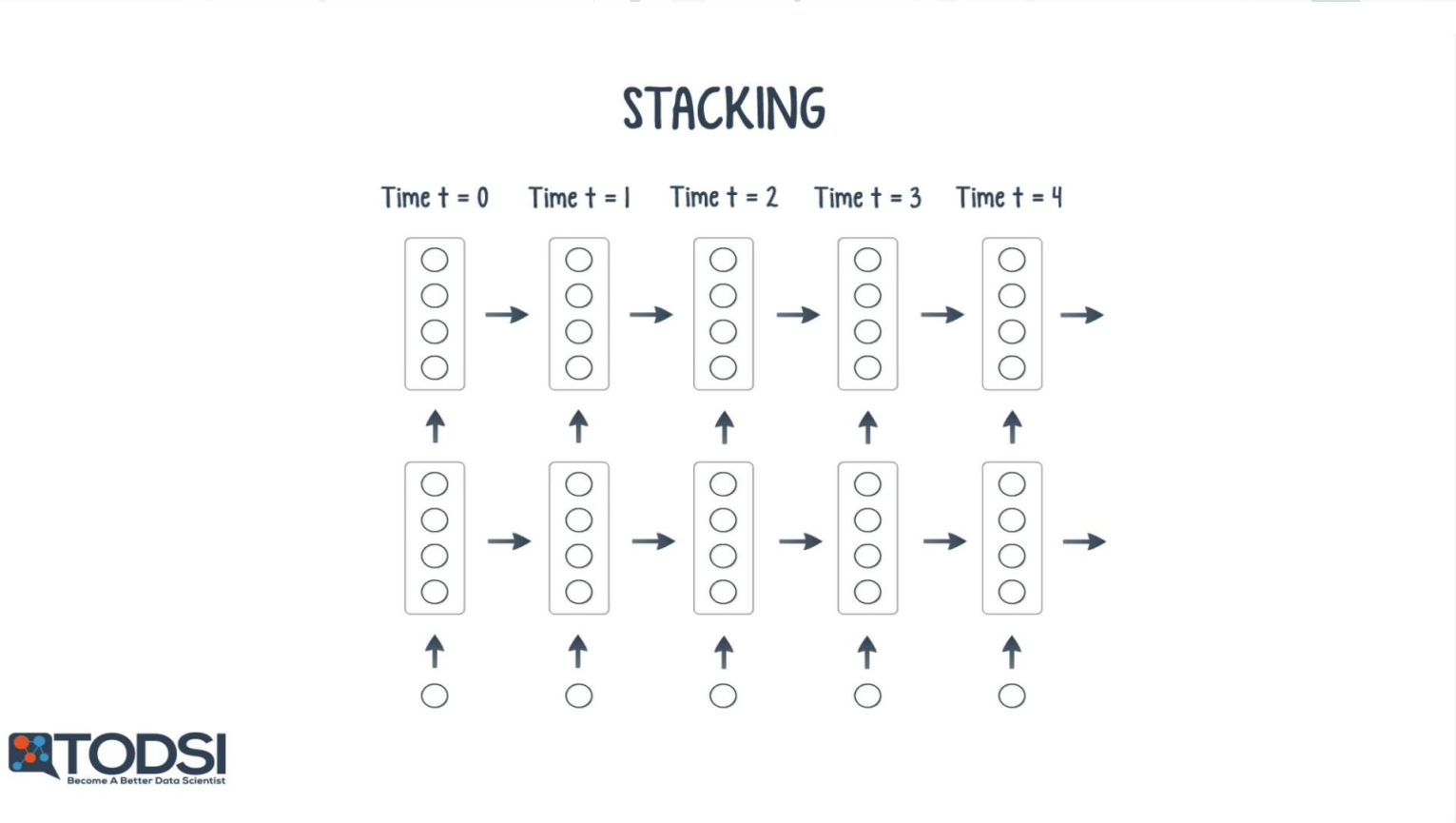
Recurrent neural network(RNN) is a class of artificial neural network where connections between units form a directed cycle. This allows it to exhibit dynamic temporal behavior. Unlike feedforward neural networks, RNNs can use their internal memory to process arbitrary sequences of inputs. This makes them applicable to tasks such as unsegmented, connected handwriting recognition or speech recognition.

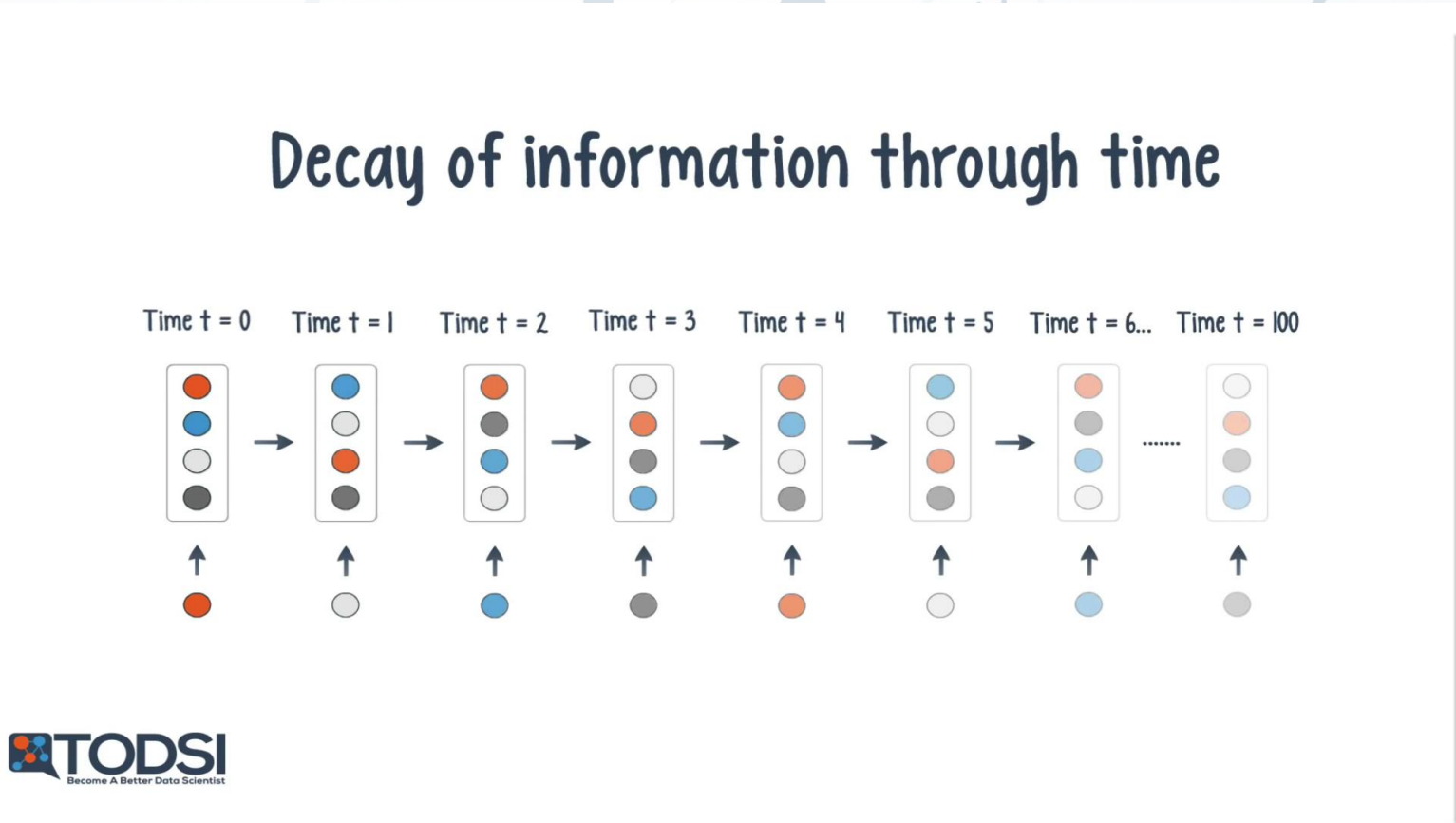
Structure of RNN



Structure of RNN



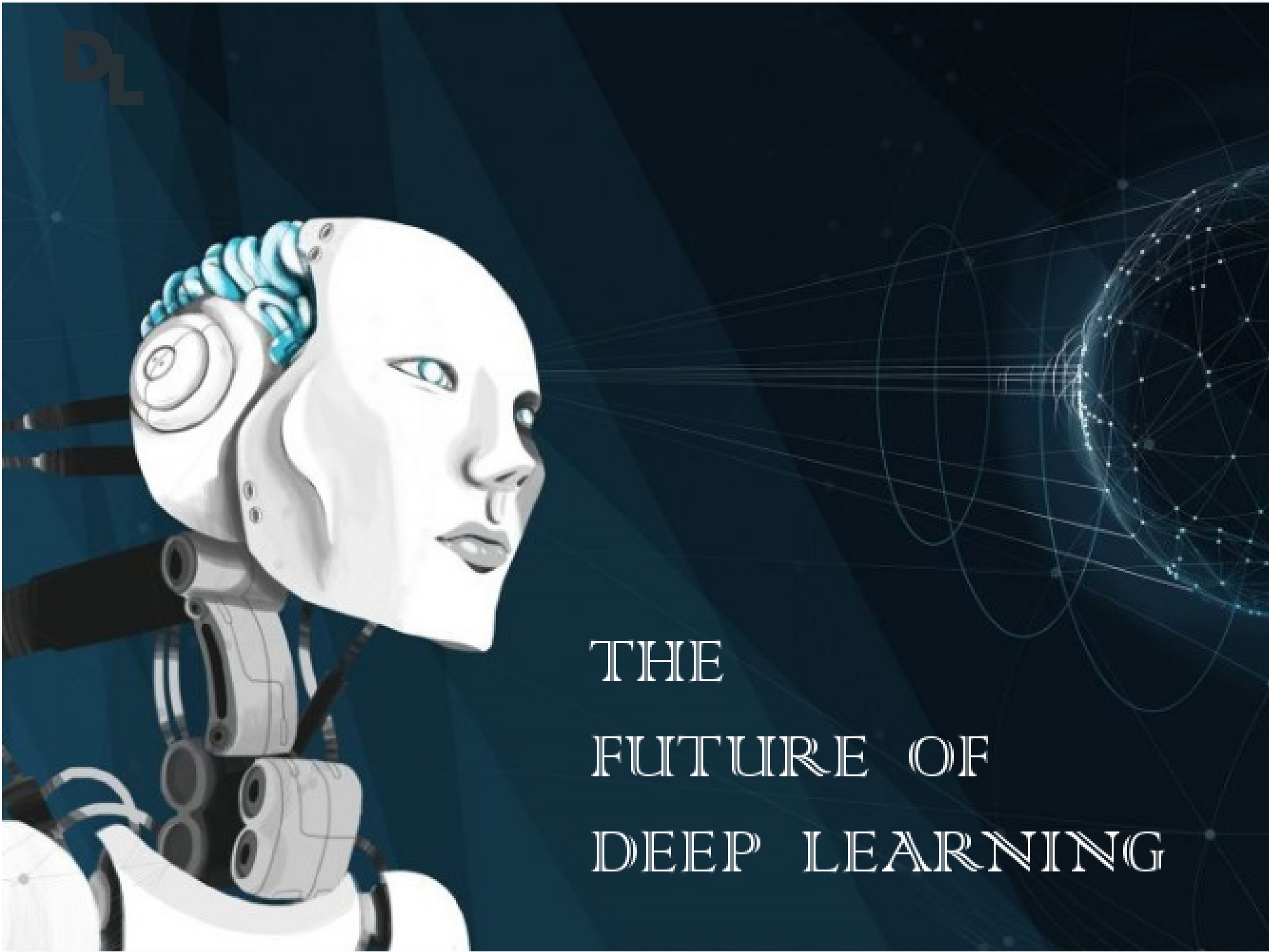




Any solution for time decay??? *Solution:* **Gating Units for e.g. LSTM**

Definition

LSTM i.e.**Long-Short Term Memory** 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
FUTURE OF
DEEP LEARNING

Unsupervised Feature Learning seems to be the future trend. Since both neural network and data sets would grow bigger and bigger, labeling everything we observed would become unreasonable and unrealistic.

So what's in hold for the future?

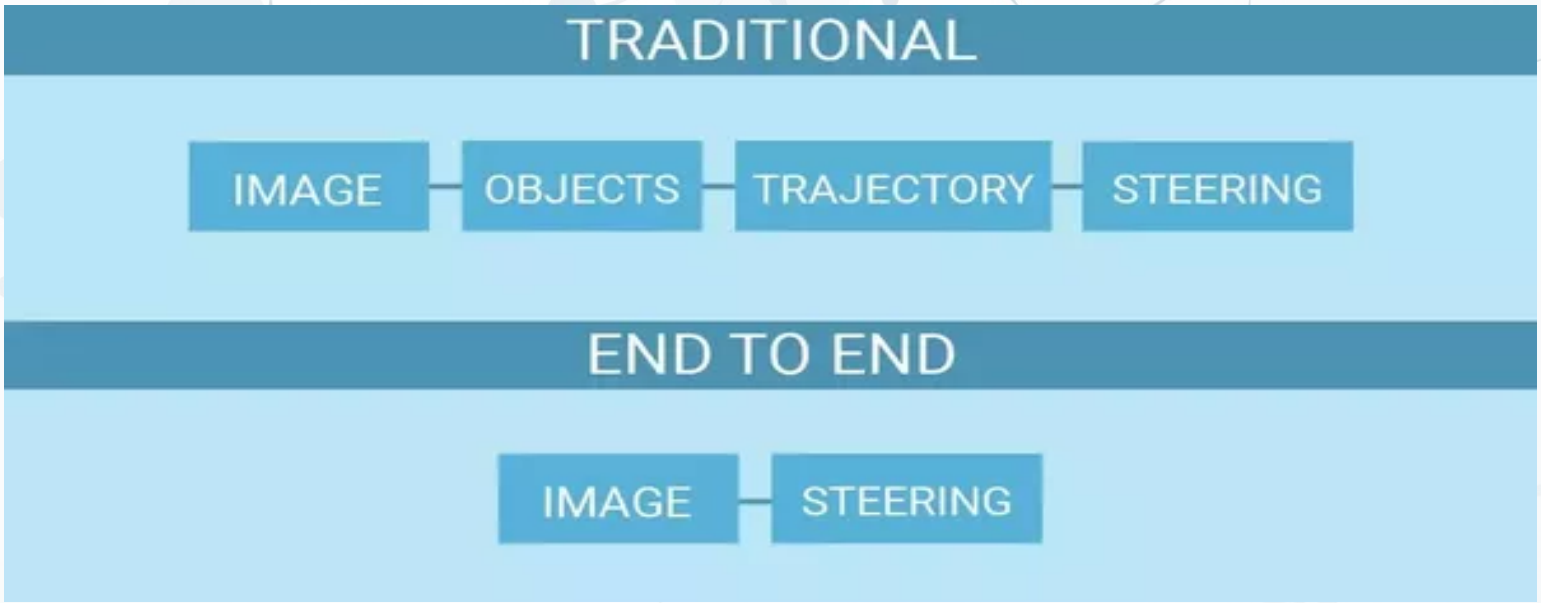
Unsupervised Feature Learning seems to be the future trend. Since both neural network and data sets would grow bigger and bigger, labeling everything we observed would become unreasonable and unrealistic.

In future, we will move away from having **"hard-coded algorithmic intelligence"** (handcrafted software) and **"learned geometric intelligence"** (deep learning). We will have instead a blend of formal *algorithmic modules that provide reasoning and abstraction capabilities*, and *geometric modules that provide informal intuition and pattern recognition capabilities*. The whole system would be learned with little or no human involvement. This would be closer to how an actual human brain works!

- **Self driving cars**

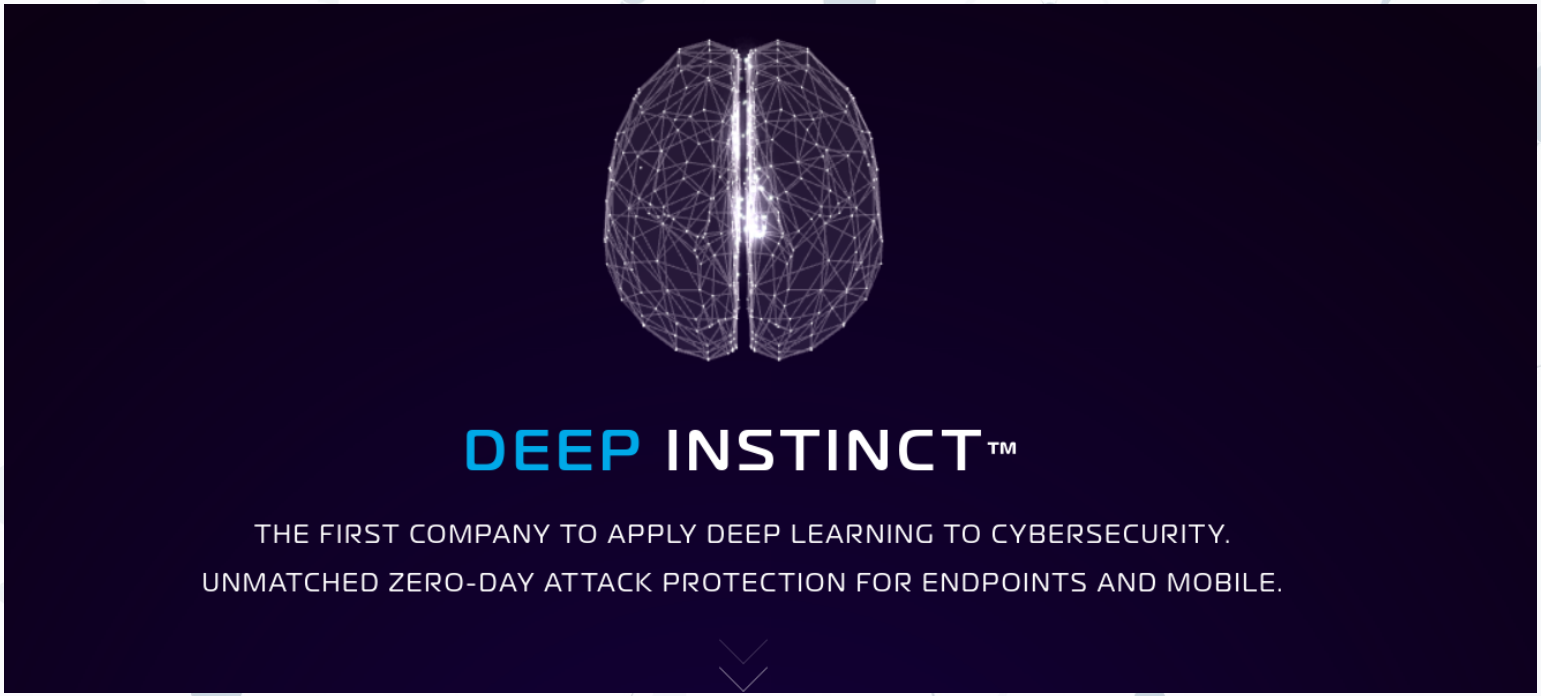
In future, end to end deep learning is likely to become more widespread as we gather more labeled data. E2E deep learning can potentially have a lot of impact for self driving cars.

The traditional approach takes an image as input, and locates the objects in the image, finds the trajectory and finally finds in what way it should steer.



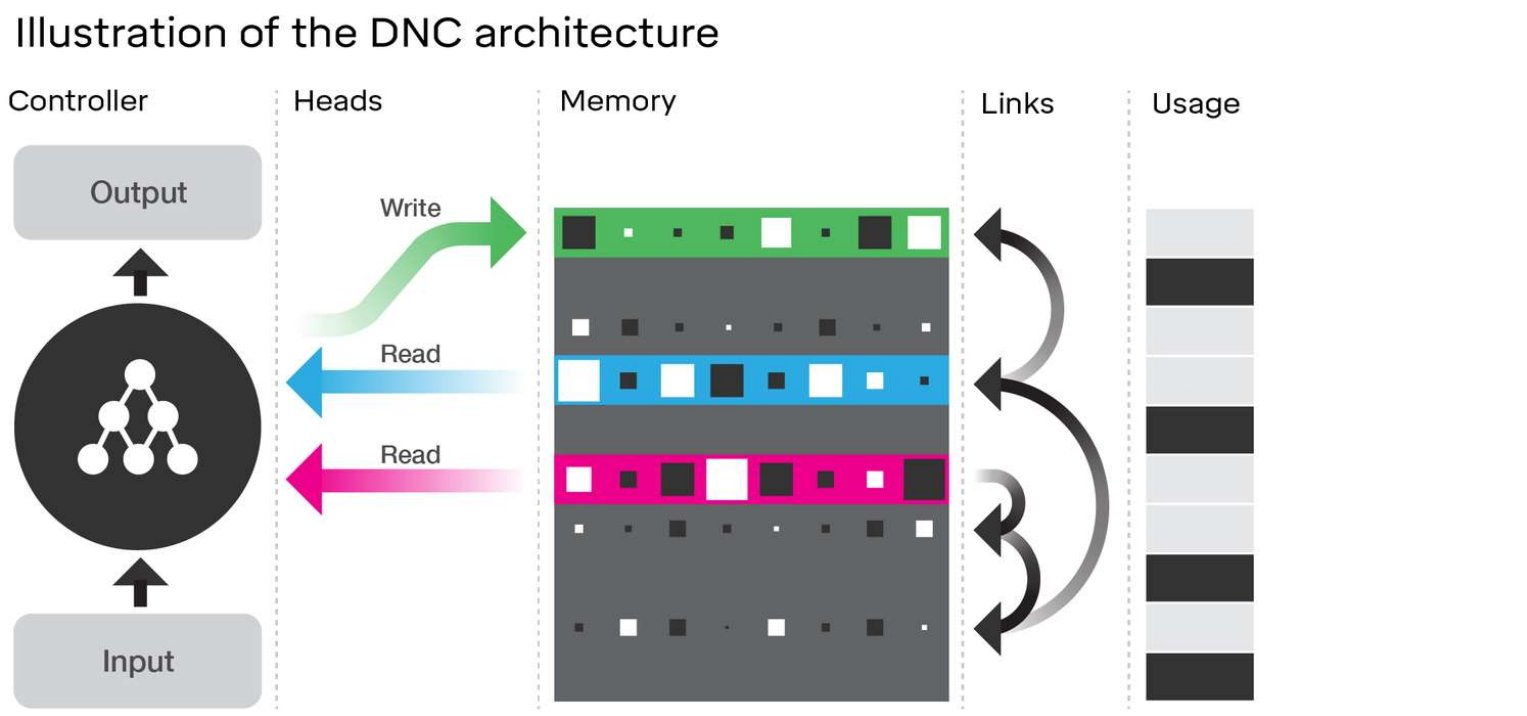
The end to end approach takes an images as input and outputs the steering.

Deep Learning in cyber-security



Deep Learning can totally change the face of cyber-security as we know it. It not only can respond to any cyber-threat faster than humans in taking pre-liminary counter measures, but it also can identify an attack or hack while it is in the process of happening by monitoring suspicious flow of data.

- **Differentiable Neural Computer**



Differentiable Neural computer or DNC is a *hybrid learning machine* *combining neural-networks with read write memory*

DNCs 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.**

The PCs and computer systems we use now have to explicitly programmed for memory management, while DNC uses previous memory usage statistics as an input in its neural frame and learns how to manage the memory as well as operations better in future.

- **Natural Language Processing- Thought Vectors**

Thought Vectors is a way of *embedding thoughts in a vector space.*

Example

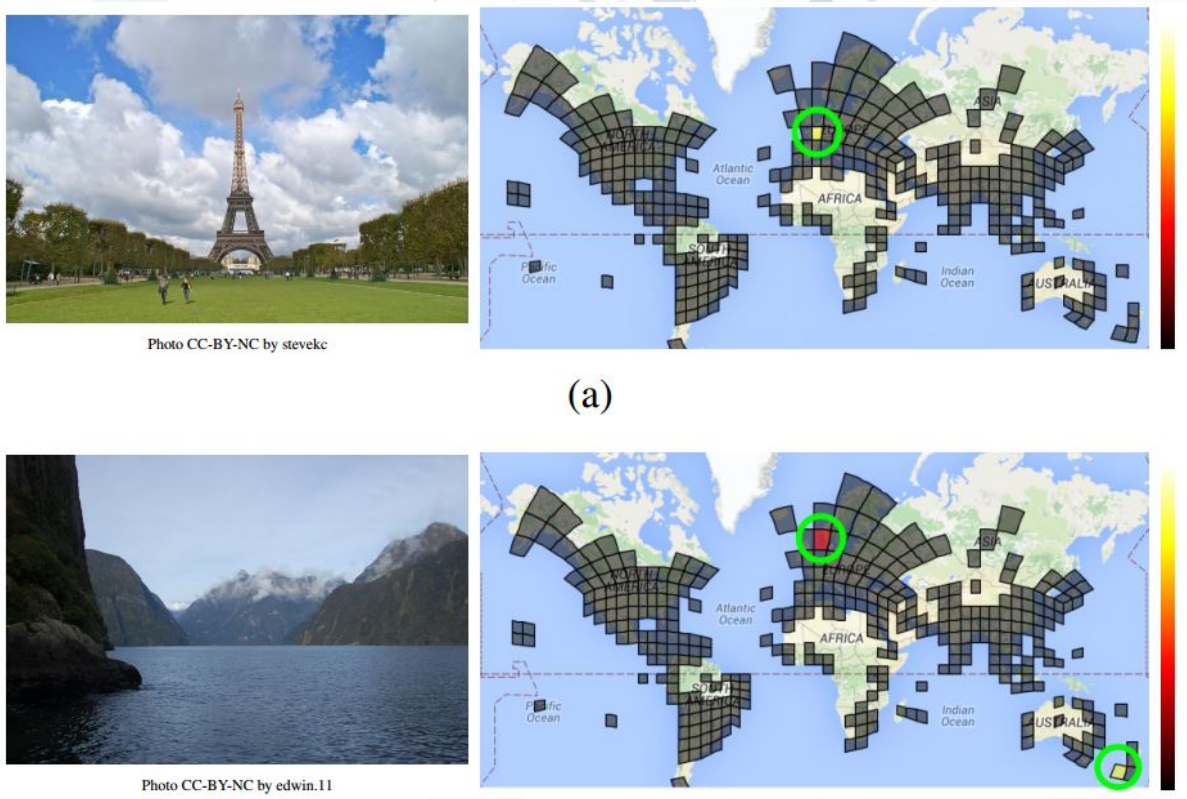
For example, **suppose a neural machine translation is trained on bilingual text.** For translation, *the input sentence (in Language 1) is first transformed into a thought vector*. This vector is used to *reconstruct the given thought* in another language (Language 2).

In future, computers would convert every sentence in a document to a thought vector, in a way that similar thoughts are nearby. Then, they would learn 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 by mimicking the thoughts expressed in the content.

Superior Image Compression

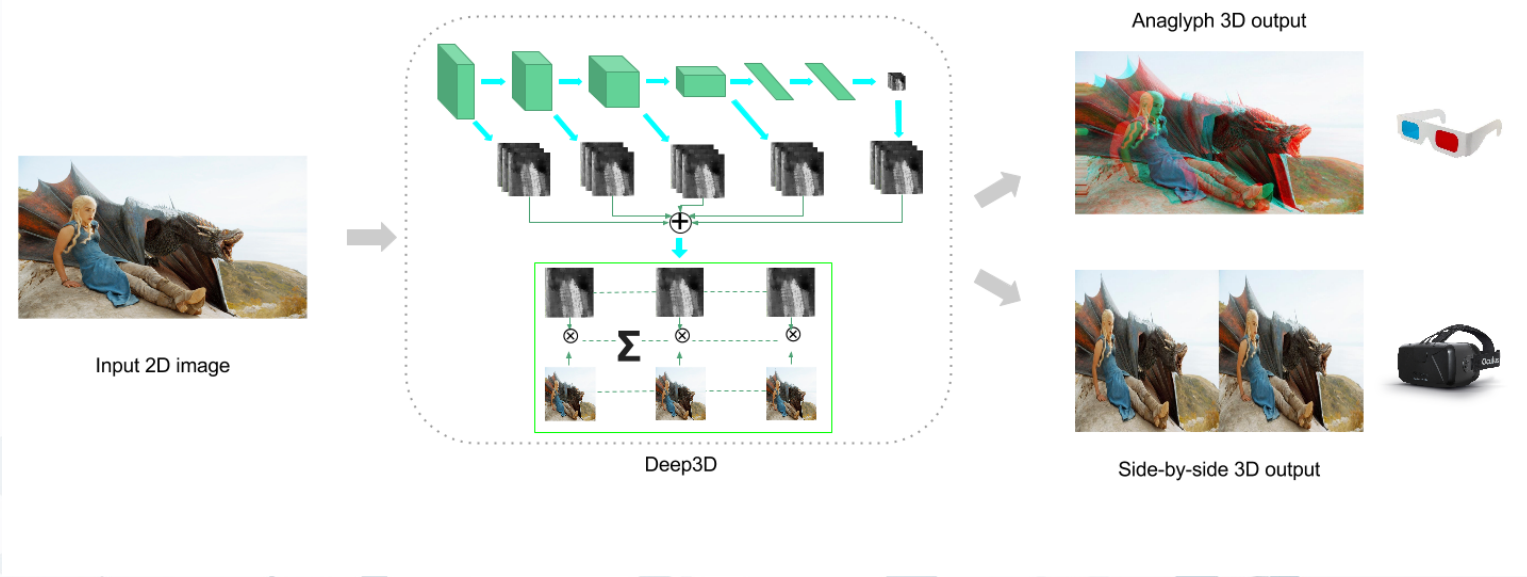
Left Original (1419KB PNG) **Center** Simple Compression (33KB JPEG) **Right** Compression via Neural Networks (24KB) i.e 25 per cent smaller for comparable image quality.

- **Image Localization- *PlaNet***



Google's PlaNet is a "Photo Geolocating Neural Network" which is able to determine the location of almost any image with superhuman ability.

Image Transformation- 2D to 3D



Deep 3D can automatically convert image/video from 2D to 3D with the help of Convolutional Neural Networks. It learns to infer 3D representations of the world based on training set of 3D movies. This can totally change the **CGI (Computer Generated Graphics)** landscape as VFX creators can be replaced by neural networks. This technique can also be extended to **VR (Virtual Reality)** as shown.

Video Sequence Prediction

PredNet is a deep convolutional recurrent network that predicts the future frames in a video sequence. These networks are able to robustly learn to predict the movement of objects.

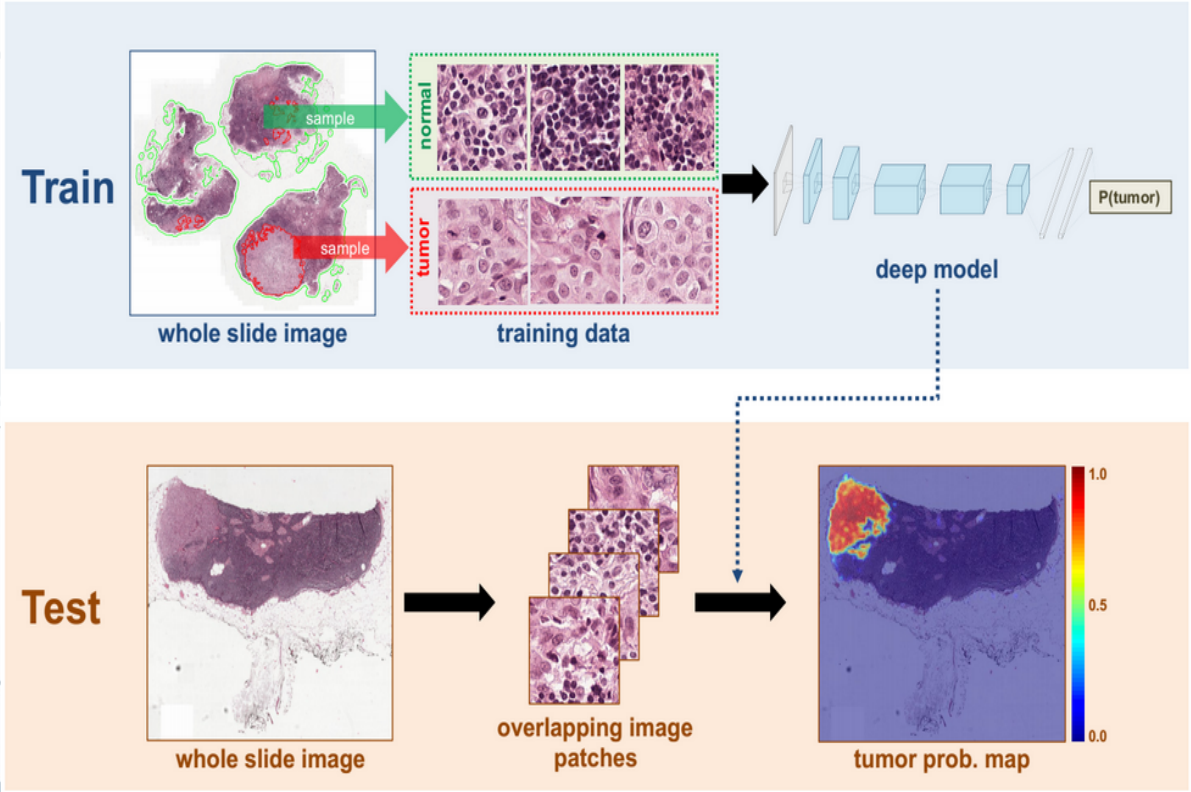
Neural Conversational Model

Human: what is the purpose of life ?	**Machine**: i 'm in the middle of nowhere .
Machine: to serve the greater good .	**Human**: what is the purpose of dying ?
Human: what is the purpose of living ?	**Machine**: to have a life
Machine: to live forever	**Human**: what is the purpose of being intelligent ?
Human: what is the purpose of existence?	**Machine**: to find out what it is.
Machine: to find out what happens when we get to the planet earth	**Human**: what is the purpose of emotions ?
Human: where are you now ?	**Machine**: I don 't know .

Philosophical Q&A – Network trained with OpenSubtitles (movie subtitles)

Neural Chatbots: Predicts the next sentence given the previous sentences in a conversation.
The LSTM (Long Short Term Memory) remembers facts, understands contexts and performs common sense reasoning in the trained domain.

Cancer Diagnoses



Deep Learning drops error rate for the cancer diagnoses by 85 per cent. Researchers trained their models with millions of labeled images to find the probablity that a patch contains cancer, eventually creating tumor probablity heatmaps.

- **Cryptography**

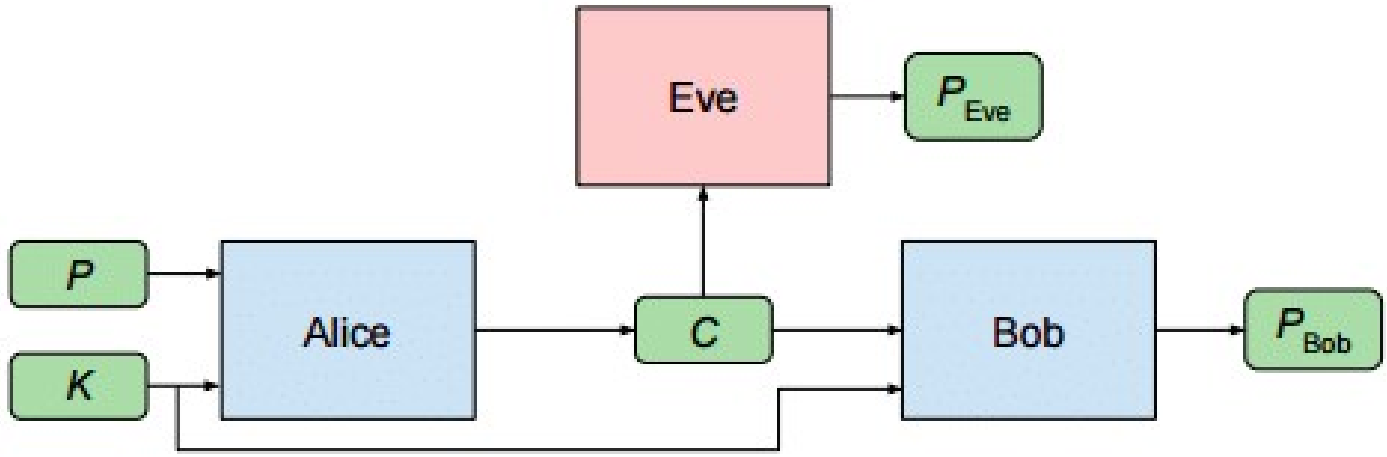
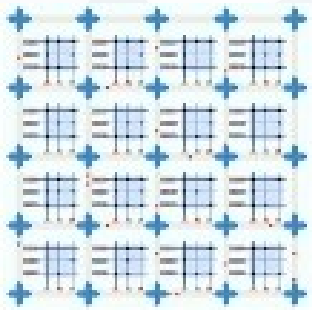


Figure 1: Alice, Bob, and Eve, with a symmetric cryptosystem.

These three are the neural networks (or neural nets) that a team from **Google Brain**, Googles research division for machine deep learning, developed to see just how well artificial intelligence (AI) can keep secrets....

The experiment involved these three, affectionately named, neural nets. The idea was for Alice and Bob to pass messages to each other while Eve tries to intercept and decode the exchanges. In order to do this, Alice and Bob were given a pre-determined set of numbers as a key to help encrypt and decrypt the message. Eve had no access to this key.
After 15,000 iterations of the scenario, Alice and Bob became adept at developing their own simple encryption technique. Alice managed to convert the original plain-text into a cipher text that Bob, in turn, was able to decrypt. Eve only managed to decipher 8 of the 16-bits in the message, where each bit was either a 1 or 0 a success rate equivalent to pure luck.

Neuromorphic Chips



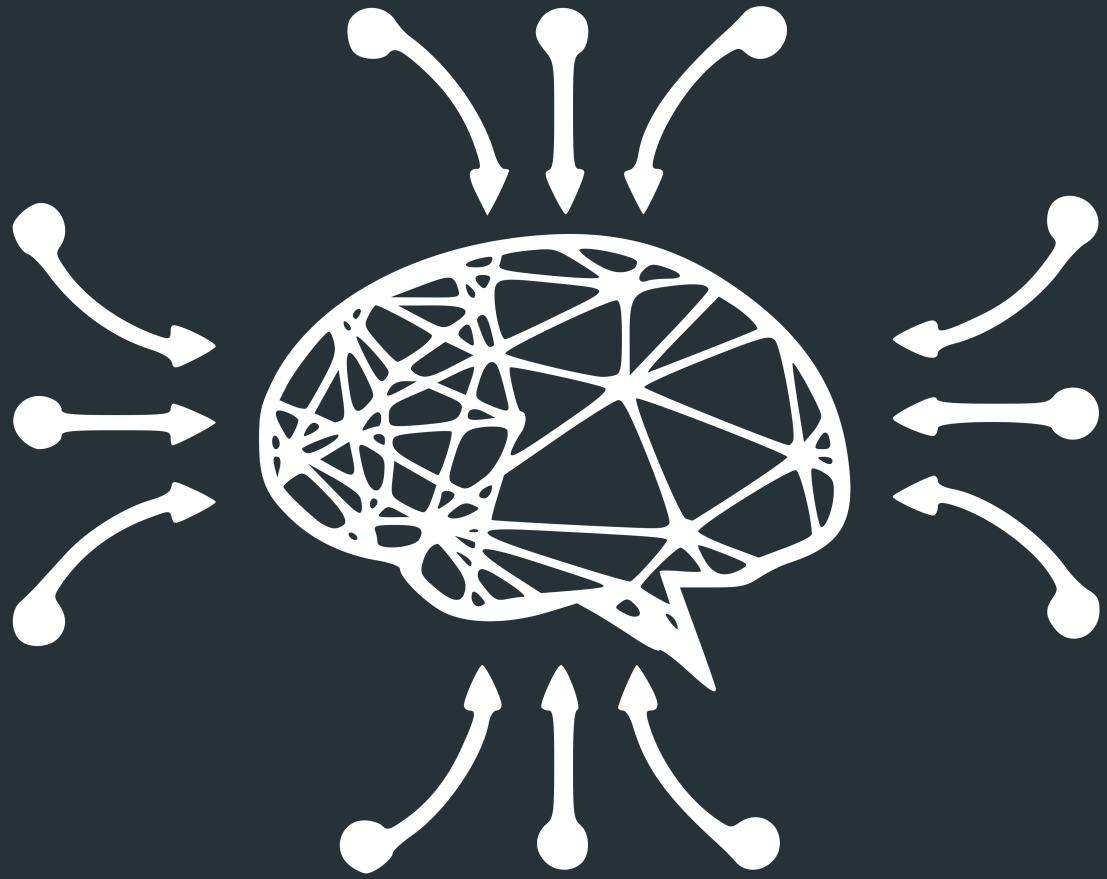
IBM TrueNorth is a brain-inspired computer chip that implements networks of integrate-and-fire spiking artificial neurons and uses only a tiny 70 mw of power –orders of magnitude less energy than traditional chips. The system is designed to be able to run deep-learning algorithms.

	What they do well	What they're good for
Neuromorphic chips	Detect and predict patterns in complex data, using relatively little electricity	Applications that are rich in visual or auditory data and that require a machine to adjust its behavior as it interacts with the world
Traditional chips (von Neumann architecture)	Reliably make precise calculations	Anything that can be reduced to a numerical problem, although more complex problems require substantial amounts of power

MIT Technology Review

- **Video Editing of the future**

- *google.com* and all its machine learning resources
- *programmableweb.com/top-10-machine-learning-apis-att-speech-ibm-watson-google-prediction*
- *adeshpande3.github.io/adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html*
- *www.analyticsvidhya.com/blog/2017/04/comparison-between-deep-learning-machine-learning*
- *dev.to/kasperfred/the-future-of-deep-learning*
- *quora.com*
- *slideshare.net/0xdata/deep-learning-through-examples*
- *appliedgo.net/perceptron*



C0nfUseD?
Go ahead, ask something

THE END

THANK YOU FOR LISTENING