Advance Analytics in a day

WELCOME

Microsoft

© 2017 Microsoft Corporation. All rights reserved.

Deep Learning

Microsoft

© 2017 Microsoft Corporation. All rights reserved.

Facundo Santiago

Cloud Solution Architect
Advance Analytics and Artificial Intelligence
CS Engineer,
MSc in Information Systems,
MSc in Business Analytics
Microsoft

medium.com/@santiagof
fasantia@microsoft.com

Microsoft

© 2017 Microsoft Corporation. All rights reserved.

Advance Analytics in a day
Roadmap

Introducción a Machine Learning .1

Presentaremos una introducción a la teoría que está detrás de Machine Learning, sus conceptos y fundamentos, cuáles son los requerimientos para emplearla y construiremos finalmente nuestro primer pipeline básico de Machine Learning utilizando los conceptos que se presentaron.

2. Modelado predictivo: clasificación y regresión

Nos adentraremos en lo que implica realizar un modelo predictivo, desde la adquisición de los datos, su transformación y preparación, elección de un algoritmo de aprendizaje y su posterior publicación para que sea consumido. Aprenderemos como hacer que nuestro modelo sea consumible a través de servicios web tipo REST y como podemos hacer uso del mismo de forma rápida incluso desde Excel, sin escribir una sola línea de código.

Análisis predictivo en tiempo real .3

Daremos nuestros primeros pasos en la implementación de un modelo de Machine Learning que sea consumido en tiempo real. Este modelo utilizará los datos que se ingestan en tiempo real y realizará predicciones como por ejemplo detectar anomalías o predecir un resultado puntual. Este tipo de despliegues se los conoce también como hot-path ya que los datos son procesados tan rápido como son generados. Implementaremos un caso real de una solución de análisis de sentimiento de los Tweets que incluyen un hashtag que nos es de interés..

4. Manipulación de datos utilizando Python y Jupyter Notebooks

Comenzaremos a familiarizarlos con el lenguaje de programación Python. Utilizaremos un SDK para manipular y transformar los datos que tenemos antes de proveerlos a nuestros modelos de Machine Learning. Una de las tareas más comunes en Data Science. Nos familiarizaremos también con una herramienta clave para cualquier científico de datos: Jupyter Notebooks y veremos como podemos correrlos fácilmente en Azure..

Natural Language Processing con servicios cognitivos .5

Microsoft Azure provee un conjunto de modelos de Machine Learning listos para utilizar empaquetados como servicios REST que cualquier desarrollador puede utilizar rápidamente. En esta sesión aprenderemos sobre estos servicios, como consumirlos, y como construir aplicaciones inteli[illegible]e sin necesidad de construir nuestro propio modelo..

Lo movimos a la sesión 8

6. Introducción a Deep Learning

Presentaremos uno de los conceptos más utilizados en Machine Leraning: Deep Learning. Se trata de una técnica de Machine Learning en donde nuestros modelos son entrenados utilizando datos provenientes de un espacio altamente dimensional (veremos también que significa esto). Implementaremos un modelo de Machine Learning para Computer Vision en Python utilizando una de las librerías mas conocidas hoy en día: PyTorch (Creada por Facebook y open-source). Veremos como podemos realizar el entrenamiento de este tipo de soluciones utilizando Azure Machine Learning Services, tanto en la nube como en nuestro propio hardware.

Operacionalizando modelos de Deep Learning .7

Veremos lo que implica llevar un modelo de Machine Learning a producción. Veremos como podemos medir la performance de nuestros modelos, evaluarlos, publicarlos como servicios y monitorearlos. Veremos también como tecnologías tales como Containers y Kubernetes serán de gran utilidad al momento de operacionalizar nuestros modelos junto con sus dependencias. Implementaremos un pipeline de Machine Learning completo que incluirá versionamiento de código, de modelos y de resultados...

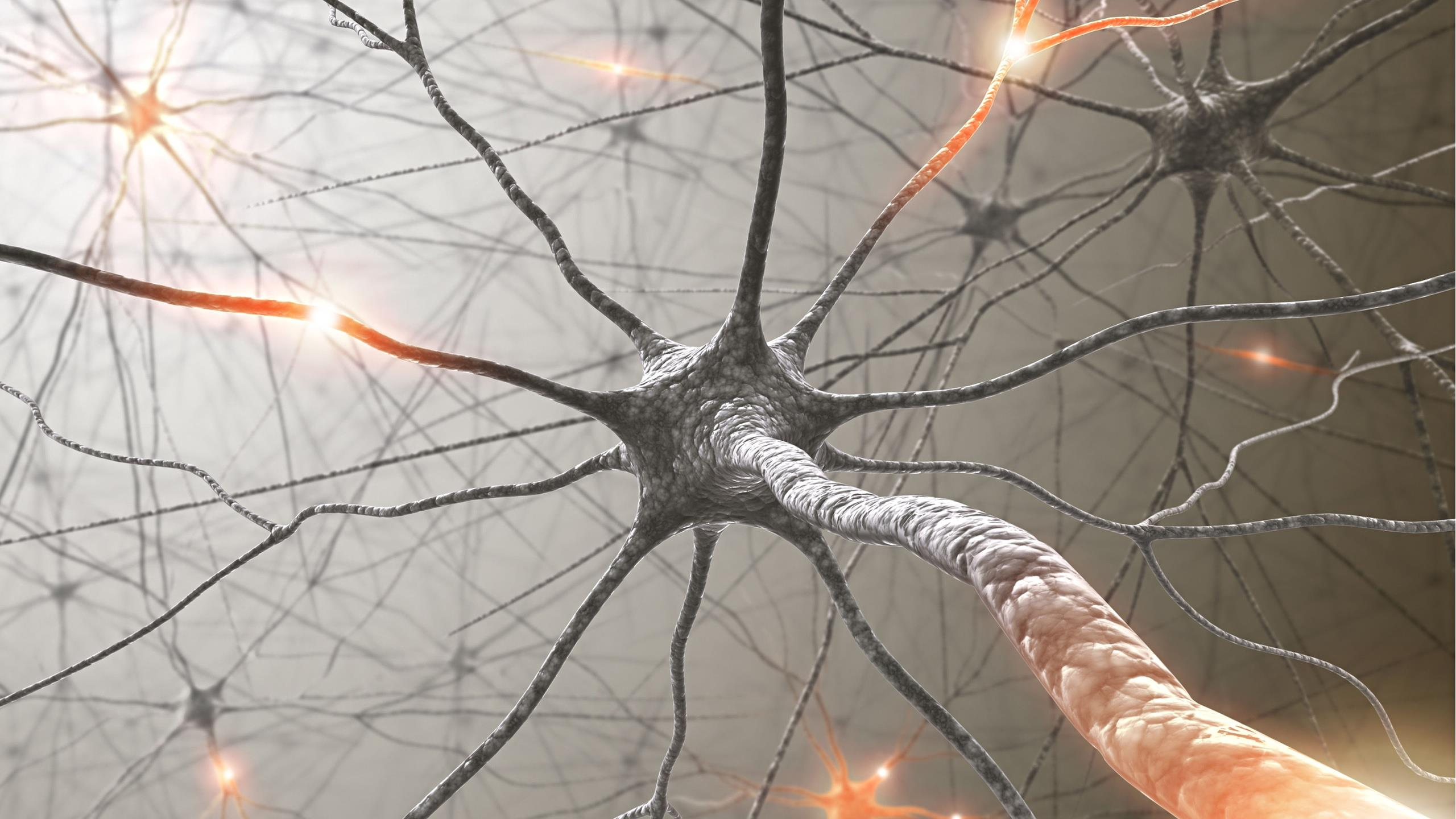
8. Procesamiento en tiempo real a escala

Avanzaremos un paso más en nuestro camino de procesamiento en tiempo real, ahora utilizando Databricks. Haremos una introducción al procesamiento distribuido utilizando clusters Spark, en este caso Databricks. Haremos doble clic en una tecnología particular conocida como structure streaming, que nos permite procesar datos en tiempo real tan pronto llegan a la plataforma

Deep Learning 101
The artificial neuron

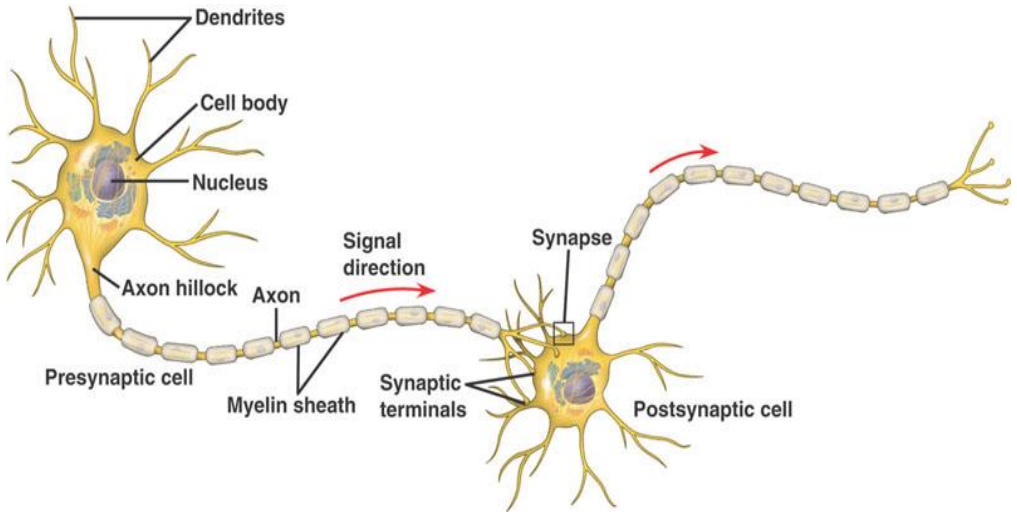
Microsoft

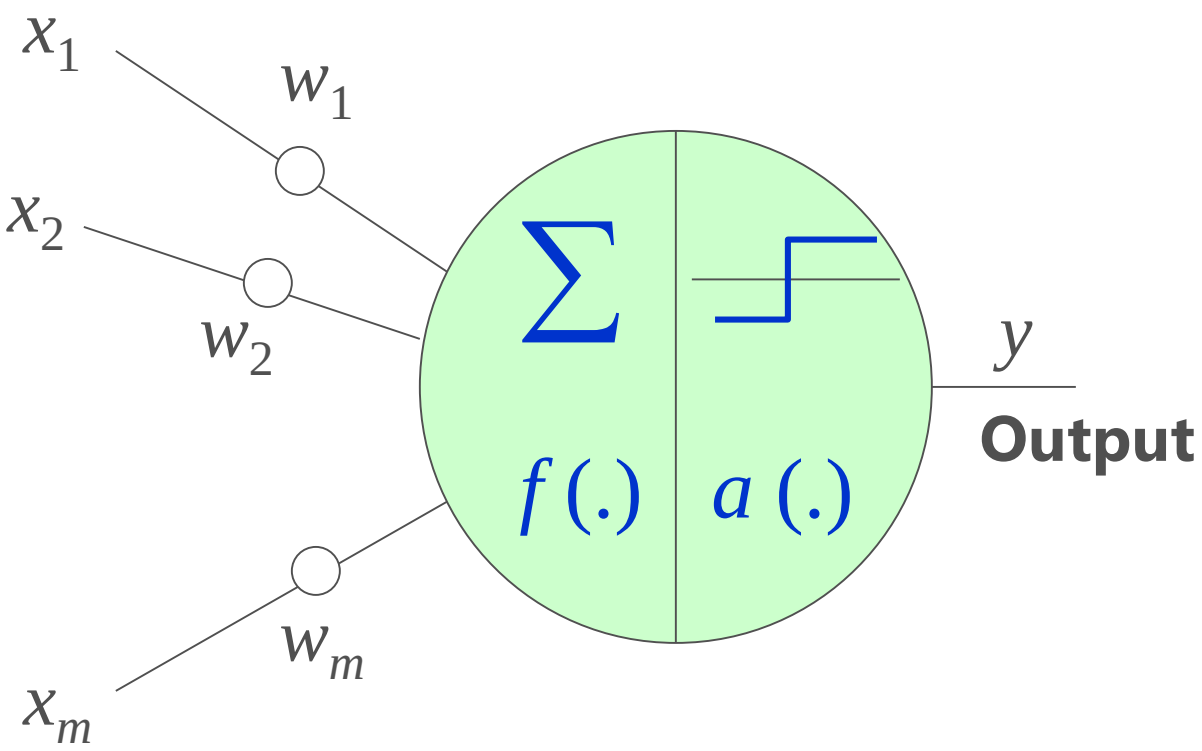
© 2017 Microsoft Corporation. All rights reserved.



The Artificial Neuron

Proposed by McCulloch and Pitts [1943] - M-P neurons





The Artificial Neuron

Integration functions and activation functions

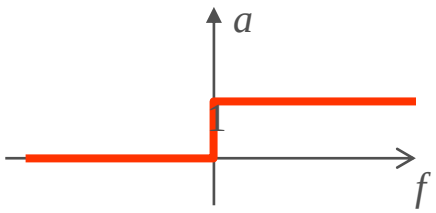
Integration functions

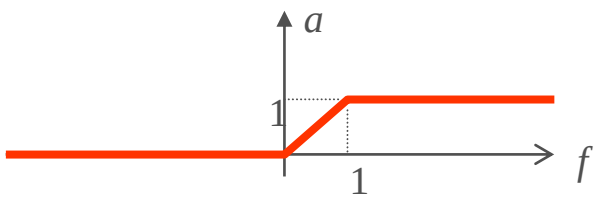
- Responsible for aggregating input signals from other neurons
- Controls boosting, inhibiting or completely suppressing the given input

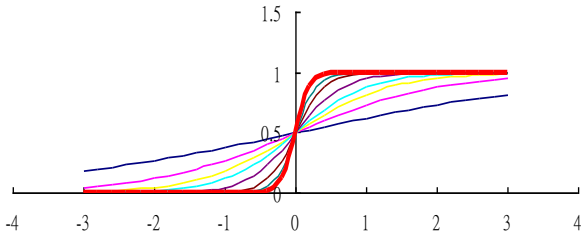
$$f(\vec{\boldsymbol{x}}) = \sum_{i=1}^{n} w_i \times x_i$$

Activation functions

- Introduces no linearity







The Artificial Neuron

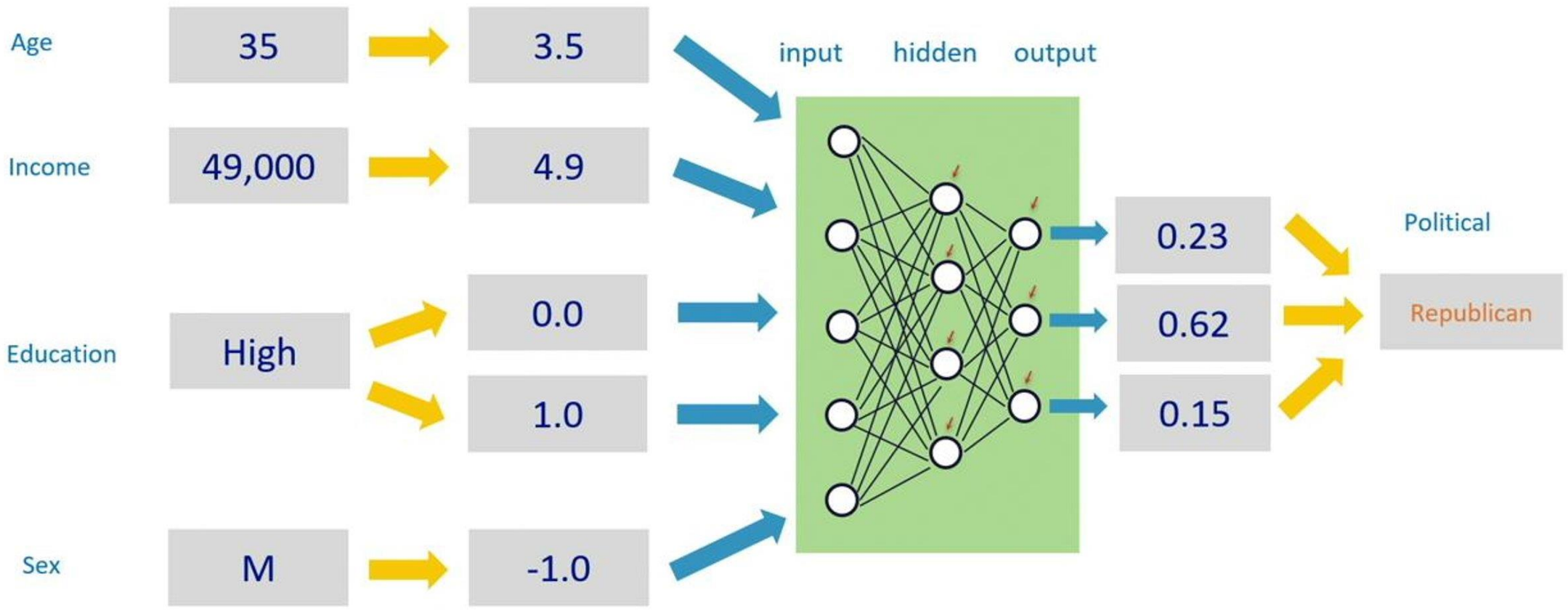
Age	Income	Education	Sex		Political
24	$24,000.00	Low	F		Democrat
62	$82,000.00	Medium	M		Republican
38	$64,000.00	High	M		Other
30	$40,000.00	Medium	F		Democrat
45	$42,000.00	High	F		Republican
35	$49,000.00	High	M		??

training data

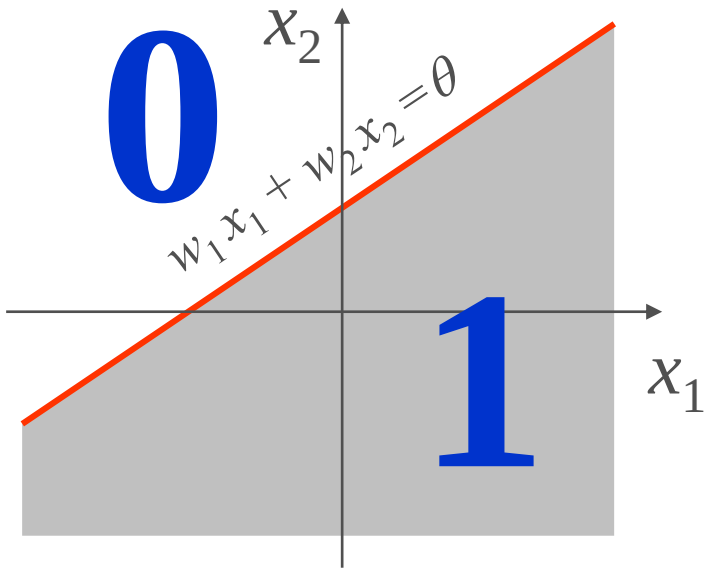
independent variables / predictors / signals / attributes / features / X-values

"the thing to classify (predict)" / label / dependent variable / Y

The Artificial Neuron



What can be achieved by single Perceptron?



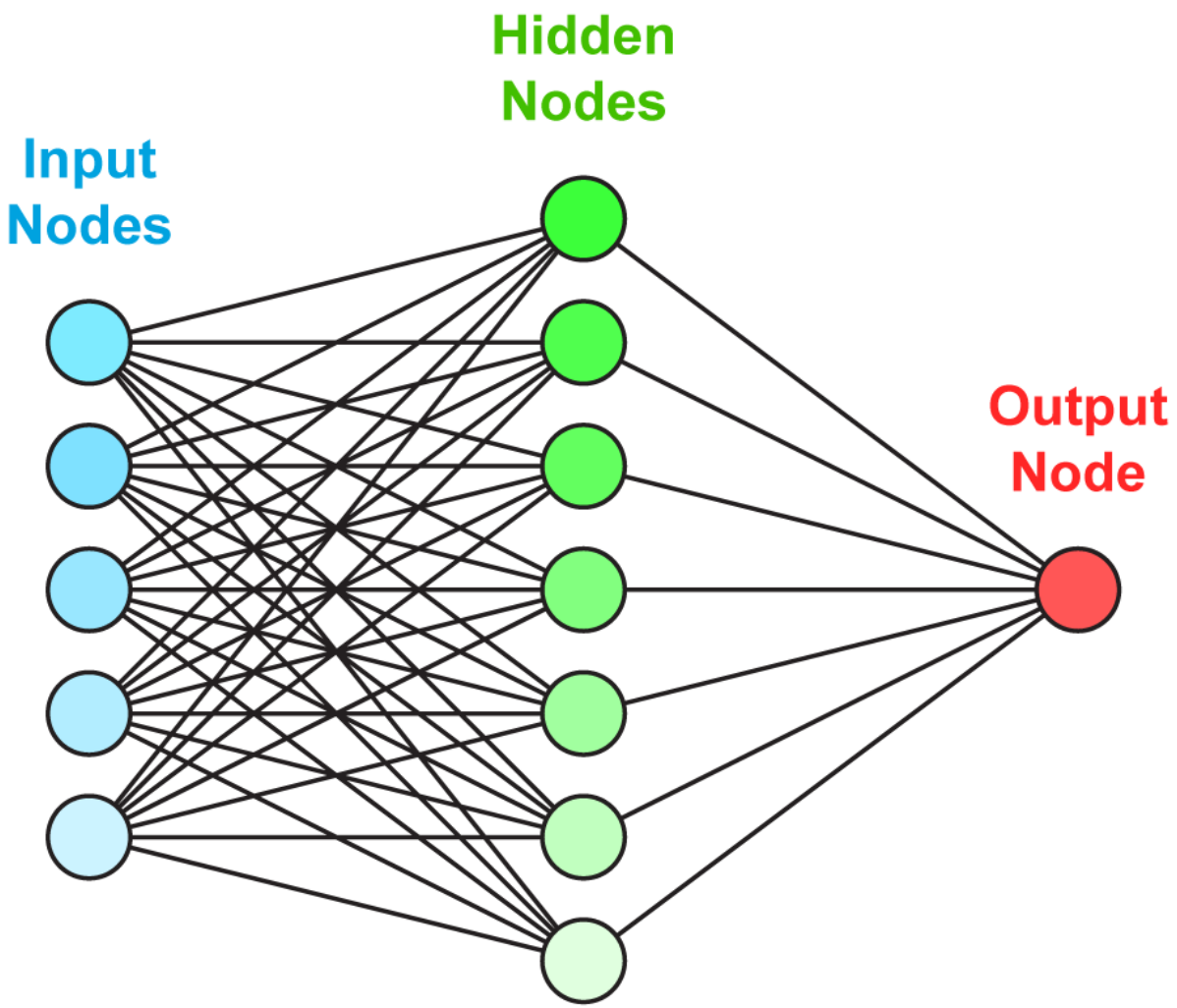
Minsky et. al. proved single neuron is no better than a linear classifier [1969]

Deep Learning 101
Neural Networks

Microsoft

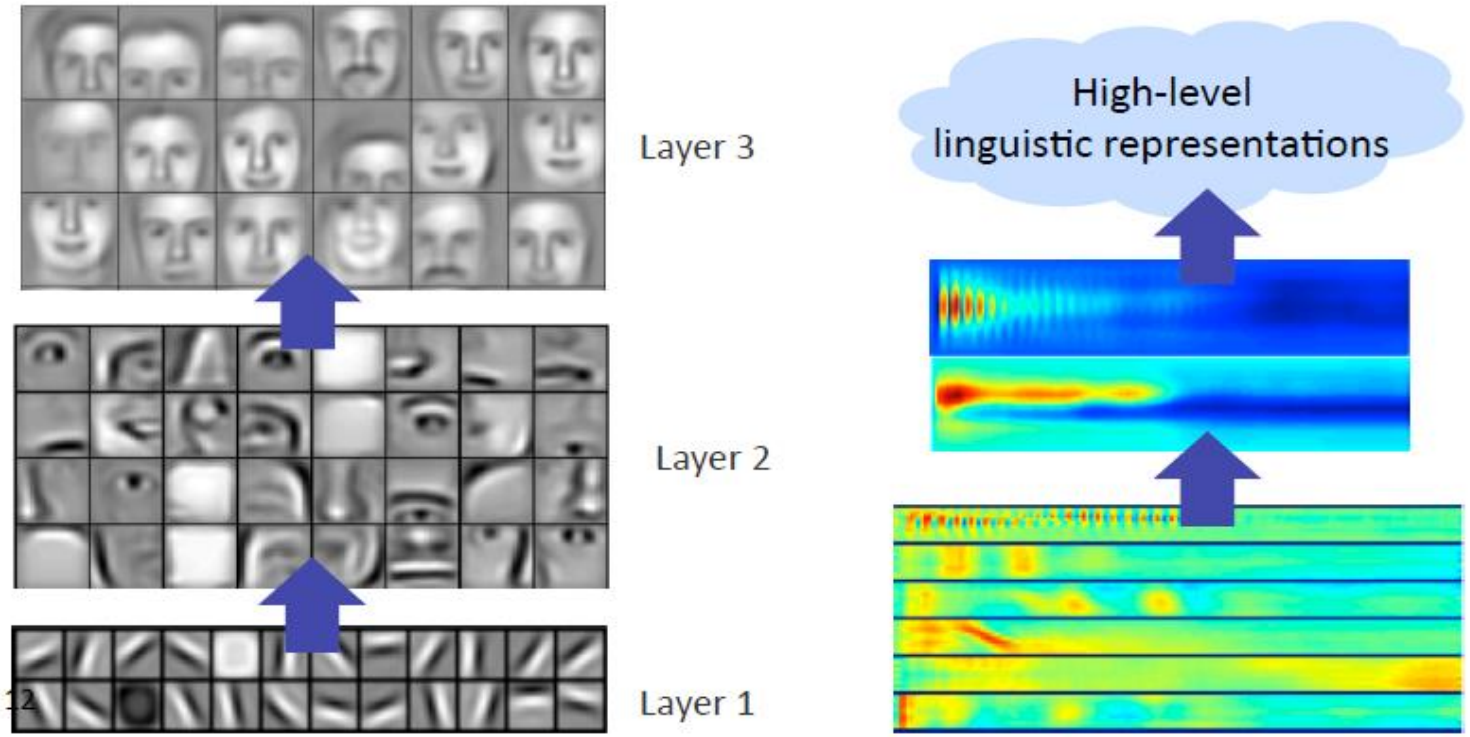
© 2017 Microsoft Corporation. All rights reserved.

MLP – Multi-Layer Perceptron



MLP – Multi-Layer Perceptron

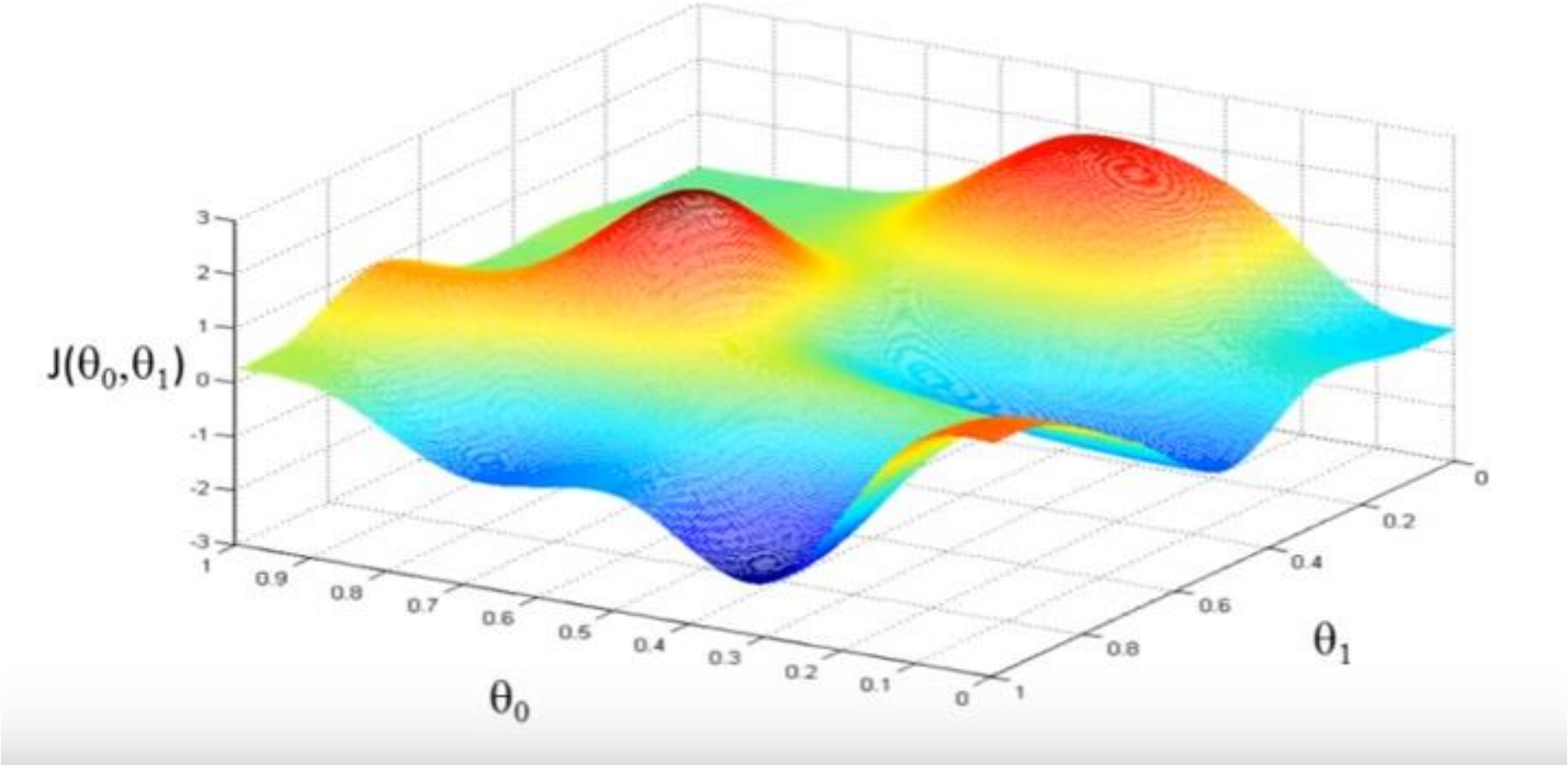
Successive model layers learn deeper intermediate representations



- Hierarchy of representations with increasing level of abstraction
- Each layer is a trainable feature transform
- Image recognition: pixel → edge → texton → motif → part → object
- Text: character → word → word group → clause → sentence → story

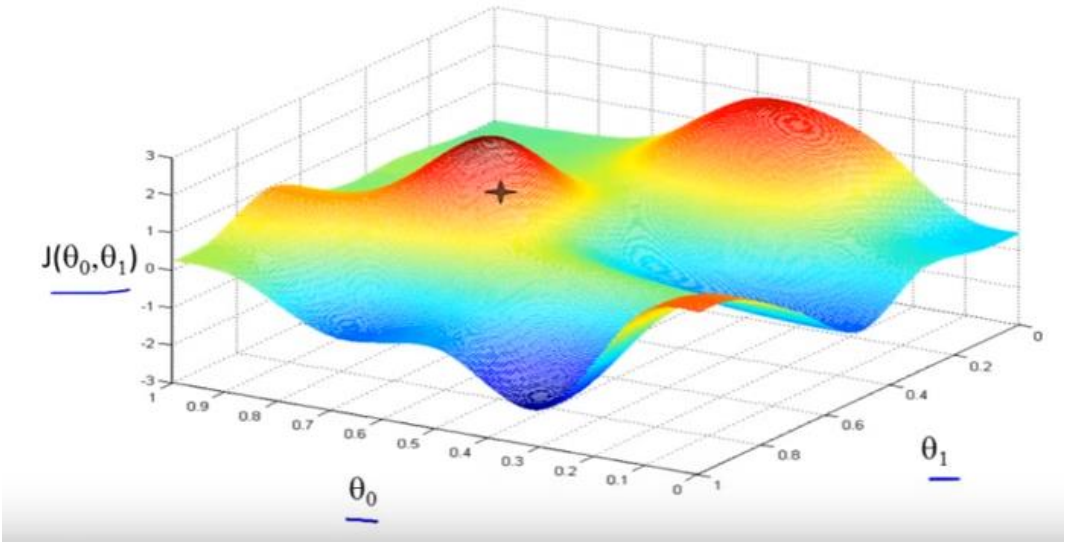
MLP – Multi-Layer Perceptron

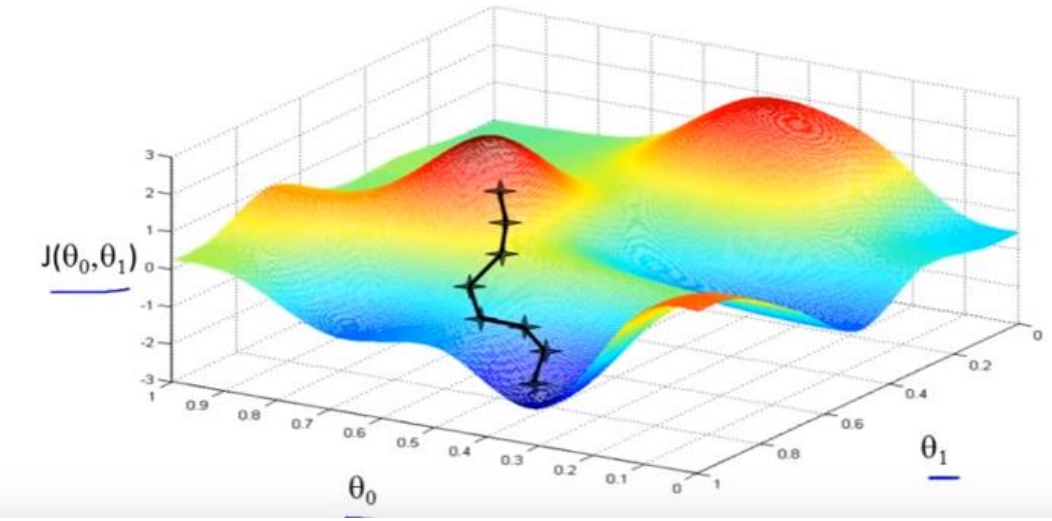
The error surface

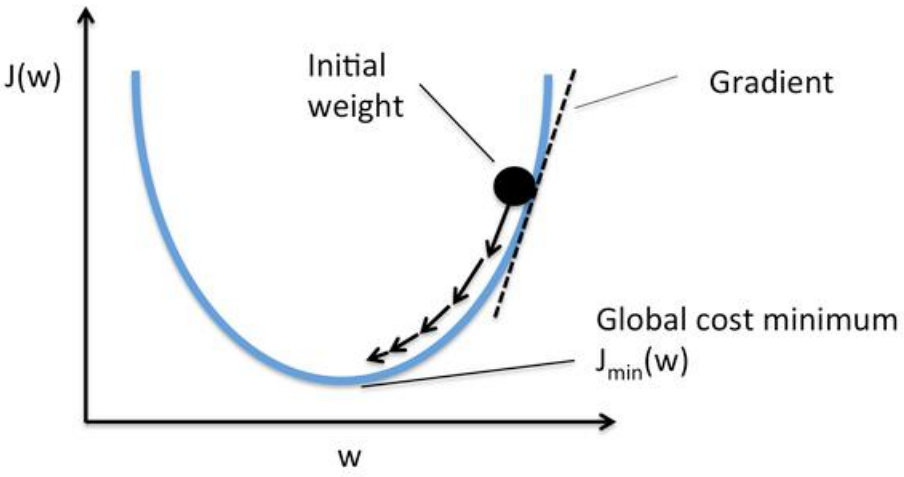


How we train this models?

By moving down the hill







Deep Learning 101
Introduction

Microsoft

© 2017 Microsoft Corporation. All rights reserved.

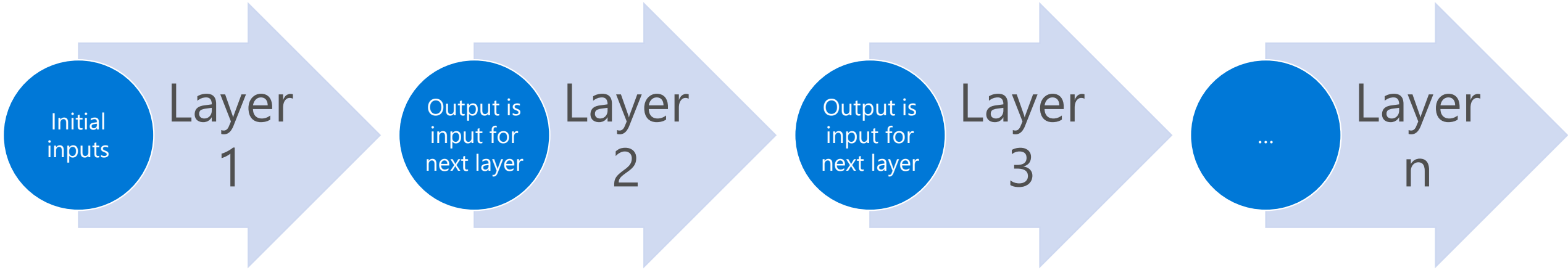
What's Deep Learning?

WIKIPEDIA
The Free Encyclopedia

Deep learning (*deep machine learning*, or *deep structured learning*, or *hierarchical learning*, or sometimes *DL*) is a branch of machine learning based on a set of algorithms that attempt to model high-level abstractions in data by using model architectures, with complex structures or otherwise, composed of multiple non-linear transformations

Deep Neural Networks (DNN)

The below applies to other algorithms as well.



Deep Learning Architectures

Supervised Learning

- CNN – Convolutional Neural Networks
- RNN – Recurrent Neural Network
- LSTM-RNN – Long-short term memory RNN
- GRD - Gated Recurrent Unit

Unsupervised Learning

- dA – Denoising autoencoders
- Stacking Autoencoders

Reinforcent Learning

- Deep Q-Networks (Q-Learning)

Deep Learning
Convolutional Neural Networks (CNN)

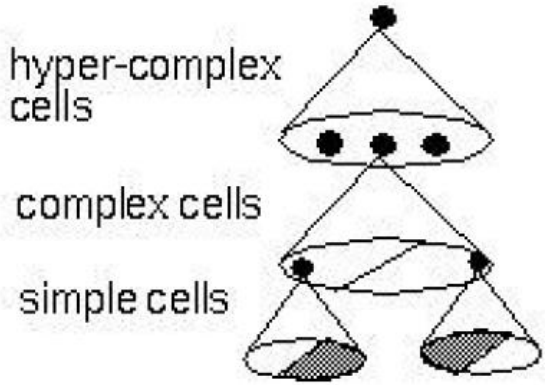
Microsoft

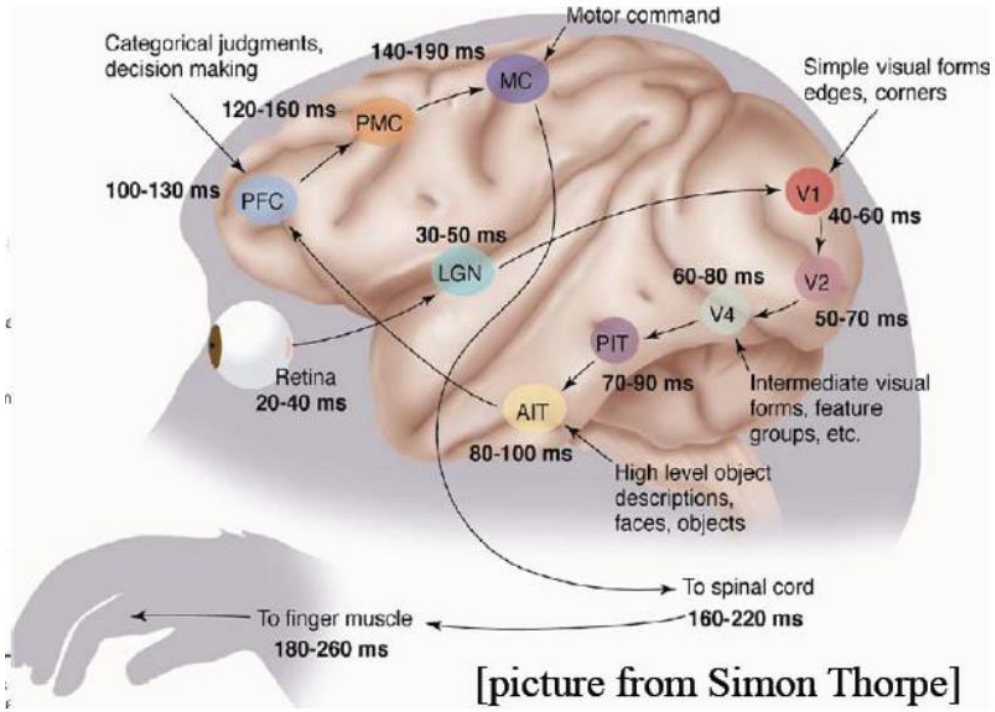
© 2017 Microsoft Corporation. All rights reserved.

Biological inspiration for CNNs

Topographical mapping in the cortex

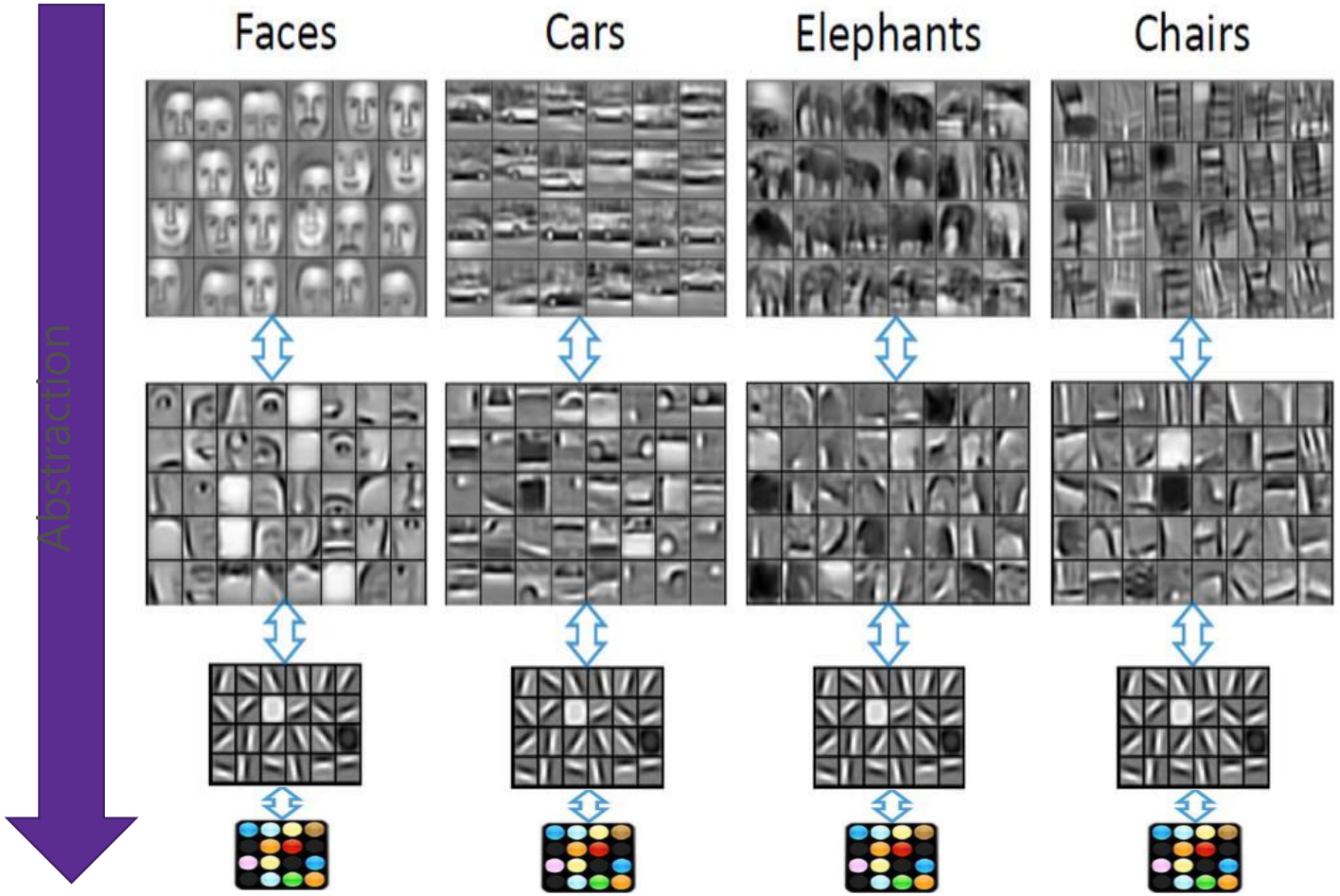
- Hubel and Weisel hierarchical mapping [1959]
  - Nearby cells in cortex represented nearby regions in the visual field
  - Visual cortex contains a complex arrangement of cells. These cells are sensitive to small sub-regions of the visual field, called a receptive field. The sub-regions are tiled to cover the entire visual field, and may overlap. These cells act as local filters over the input space and are well-suited to exploit the strong spatially local correlation present in natural images.
  - Additionally, two basic cell types have been identified: Simple cells respond maximally to specific edge-like patterns within their receptive field. Complex cells have larger receptive fields and are locally invariant to the exact position of the pattern.
- The neocognitron is a hierarchical, multilayered artificial neural network proposed by Kunihiko Fukushima in the 1980s.





[picture from Simon Thorpe]

Hierarchical Approach



VISION

pixels → edge → texton → motif → part → object

SPEECH

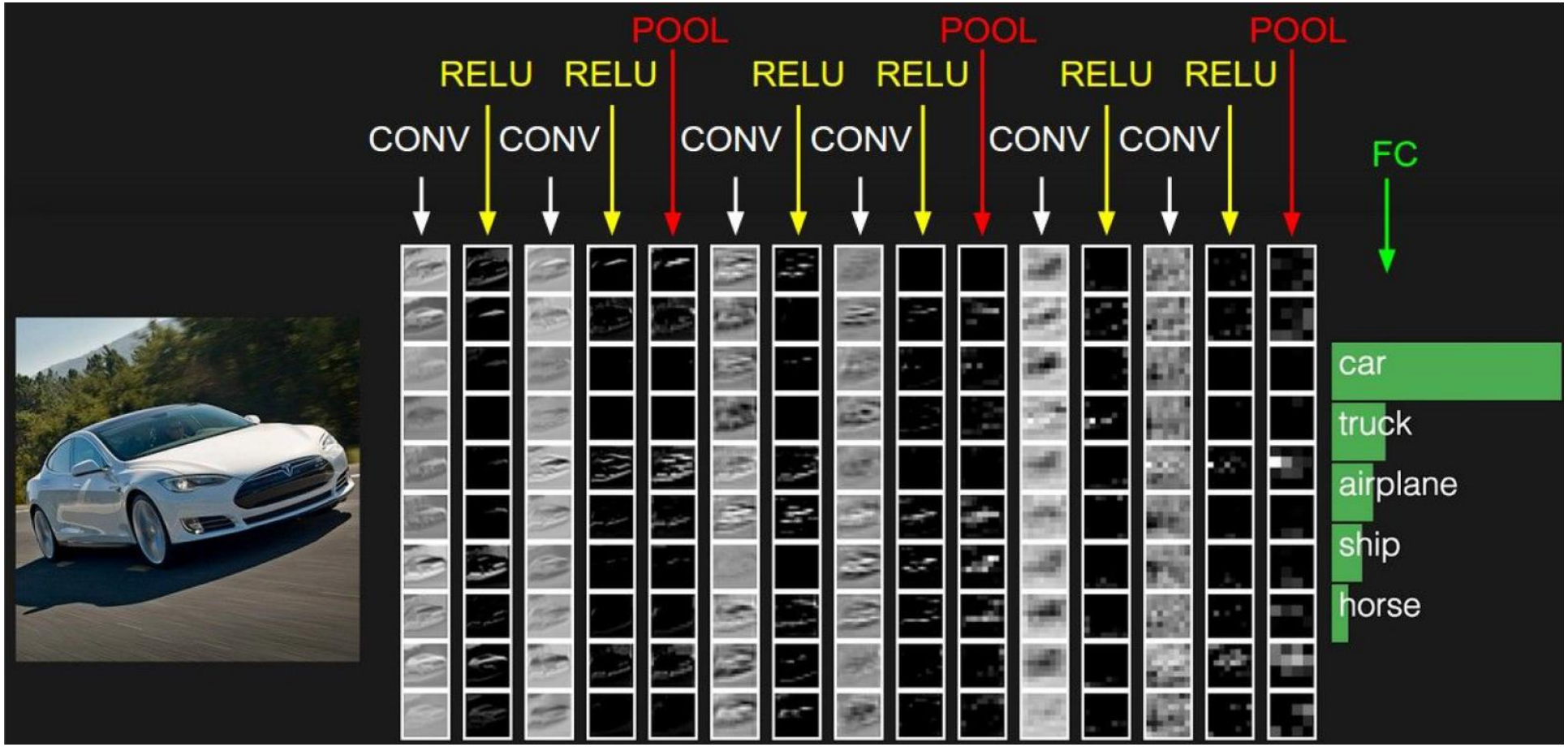
sample → spectral band → formant → motif → phone → word

NLP

character → word → NP/VP/.. → clause → sentence → story

From: Ranzanto

ConvNet with Pooling and FC Layers



Convolution Networks are everywhere now

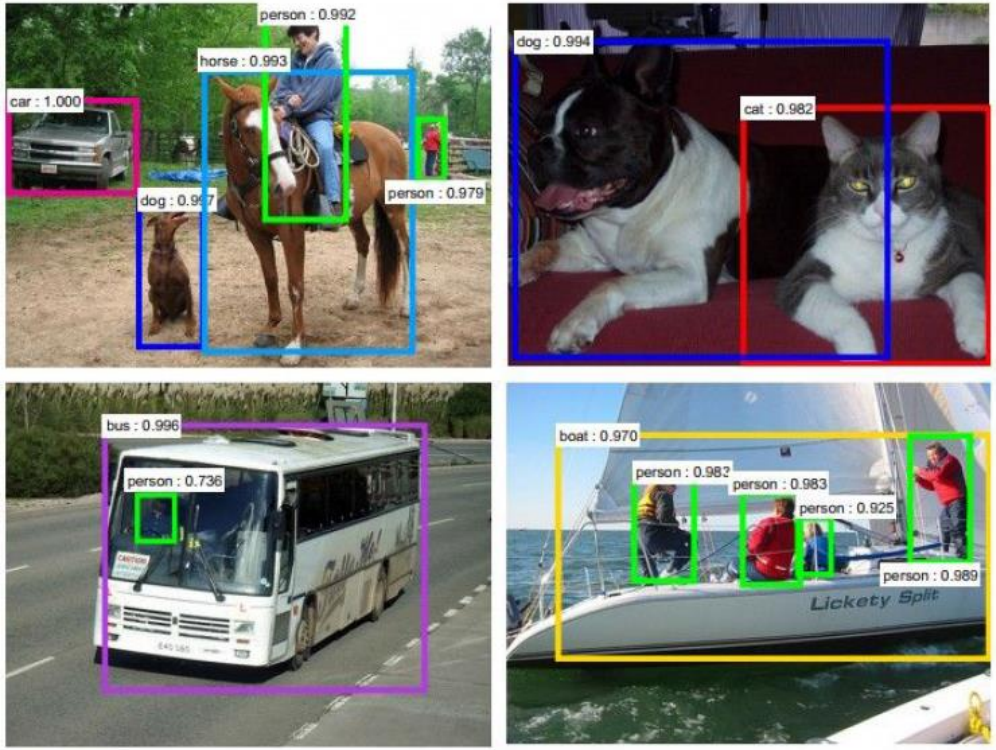
Classification

Retrieval

[Krizhevsky 2012]

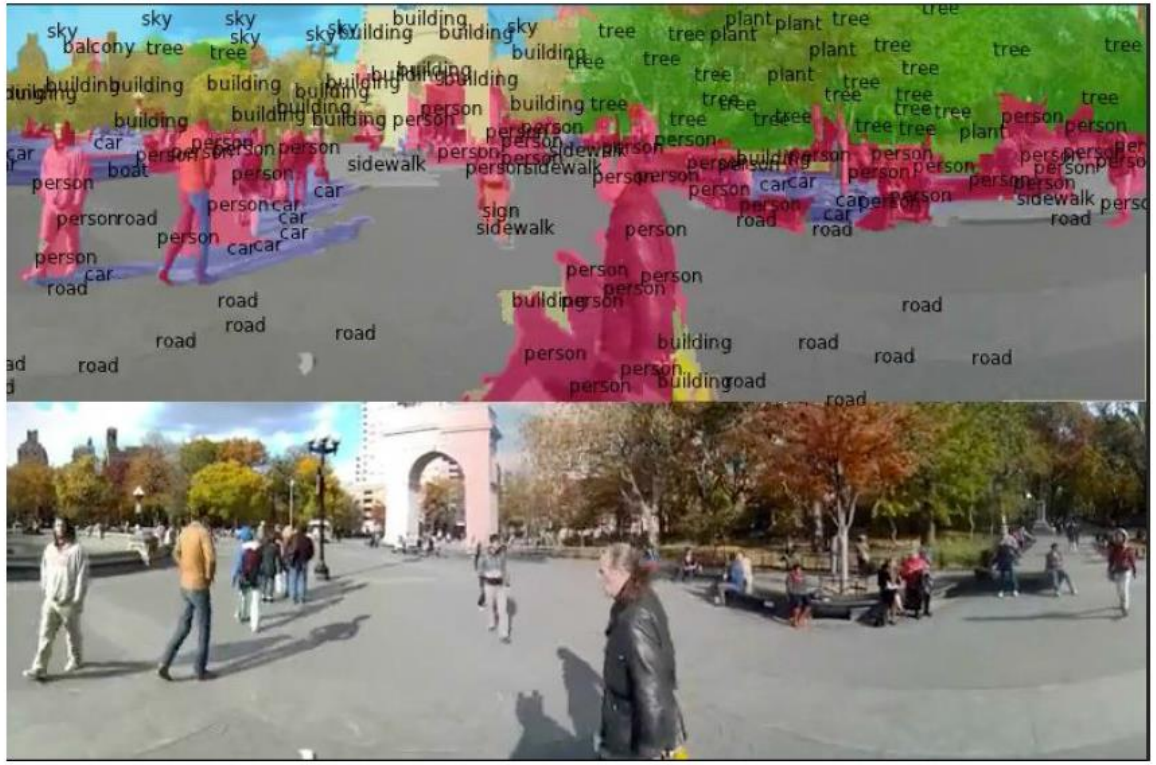
Convolution Networks are everywhere now

Detection



[Faster R-CNN: Ren, He, Girshick, Sun 2015]

Segmentation

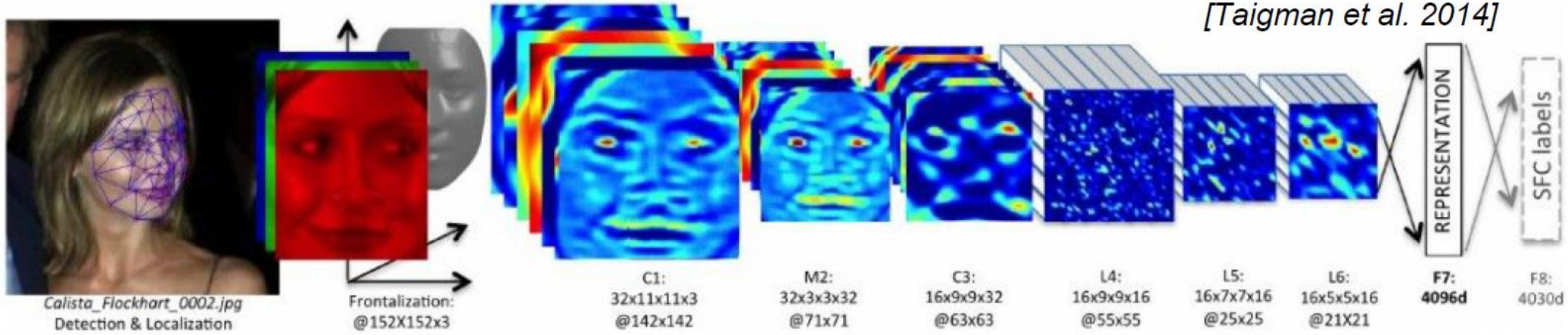


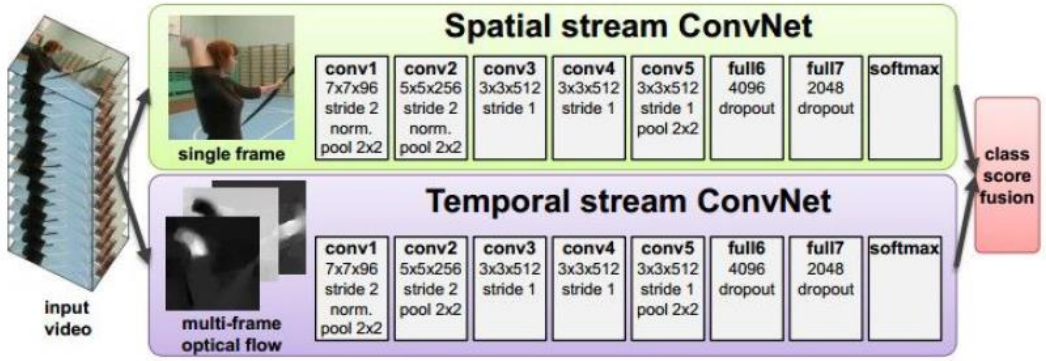
[Farabet et al., 2012]

Convolution Networks are everywhere now

self-driving cars

Convolution Networks are everywhere now

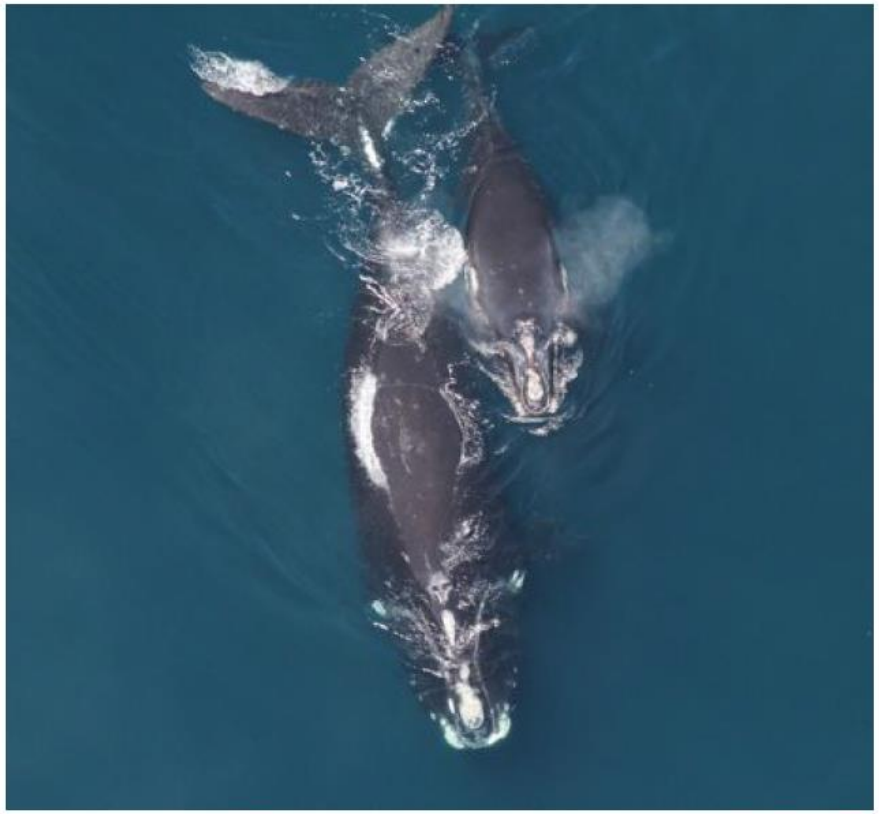




[Simonyan et al. 2014]

[Goodfellow 2014]

Convolution Networks are everywhere now



Whale recognition, Kaggle Challenge

Mnih and Hinton, 2010

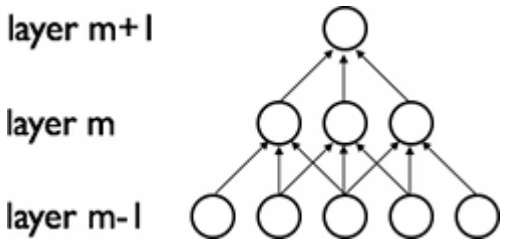
Convolution Networks are everywhere now

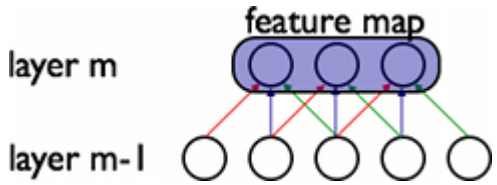
Image Captioning

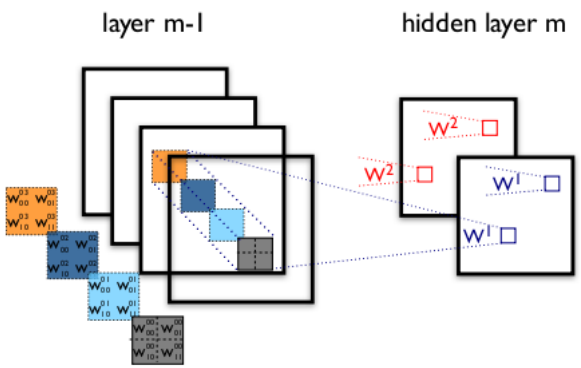
[Vinyals et al., 2015]

Convolutional Neural Networks (LeNet)

- Sparse local Connectivity between adjacent layers
- Shared Weights
  - Each filter h_i is replicated across the entire visual field.
  - Replicated units share the same parameterization (weight vector and bias) and form a feature map.
  - Weights of the same color are shared.
  - Weight sharing increases learning efficiency by greatly reducing the number of free parameters being learnt.
- To form a richer representation of the data, each hidden layer is composed of multiple feature maps, $\{h^{(k)}, k = 0..K\}$
- Max-Pooling or down-sampling or sub-sampling.
  - Max-pooling partitions the input image into a set of non-overlapping rectangles and, for each such sub-region, outputs the maximum value.
  - Max-pooling is useful in vision for two reasons:
    - By eliminating non-maximal values, it reduces computation for upper layers.
    - It provides a form of translation invariance.







ImageNet Challenge Winning Architectures

Year 2010

NEC-UIUC

Dense grid descriptor: HOG, LBP

Coding: local coordinate, super-vector

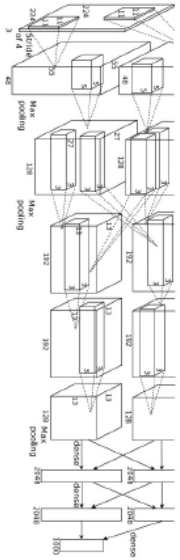
Pooling, SPM

Linear SVM

[Lin CVPR 2011]

Year 2012

SuperVision



[Krizhevsky NIPS 2012]

7 layers

Year 2014

GoogLeNet

[Szegedy arxiv 2014]

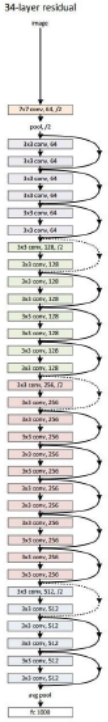
27 layers

VGG

[Simonyan arxiv 2014]

Year 2015

MSRA



151 layers

Thank you

Coming next: Operationalization of Deep Learning models

Microsoft

© 2017 Microsoft Corporation. All rights reserved.