Machine Learning

Thomas Trappenberg

Dalhousie University

Fall 2017

- Building machines that learn from examples rather than explicit programming a specific task

- Solve problems that have been difficult to solve otherwise (object recognition, text translation, speech recognition, …)

- How to build adaptive systems

- How do brains work (real intelligence)

Google

amazon

facebook

DeepMind

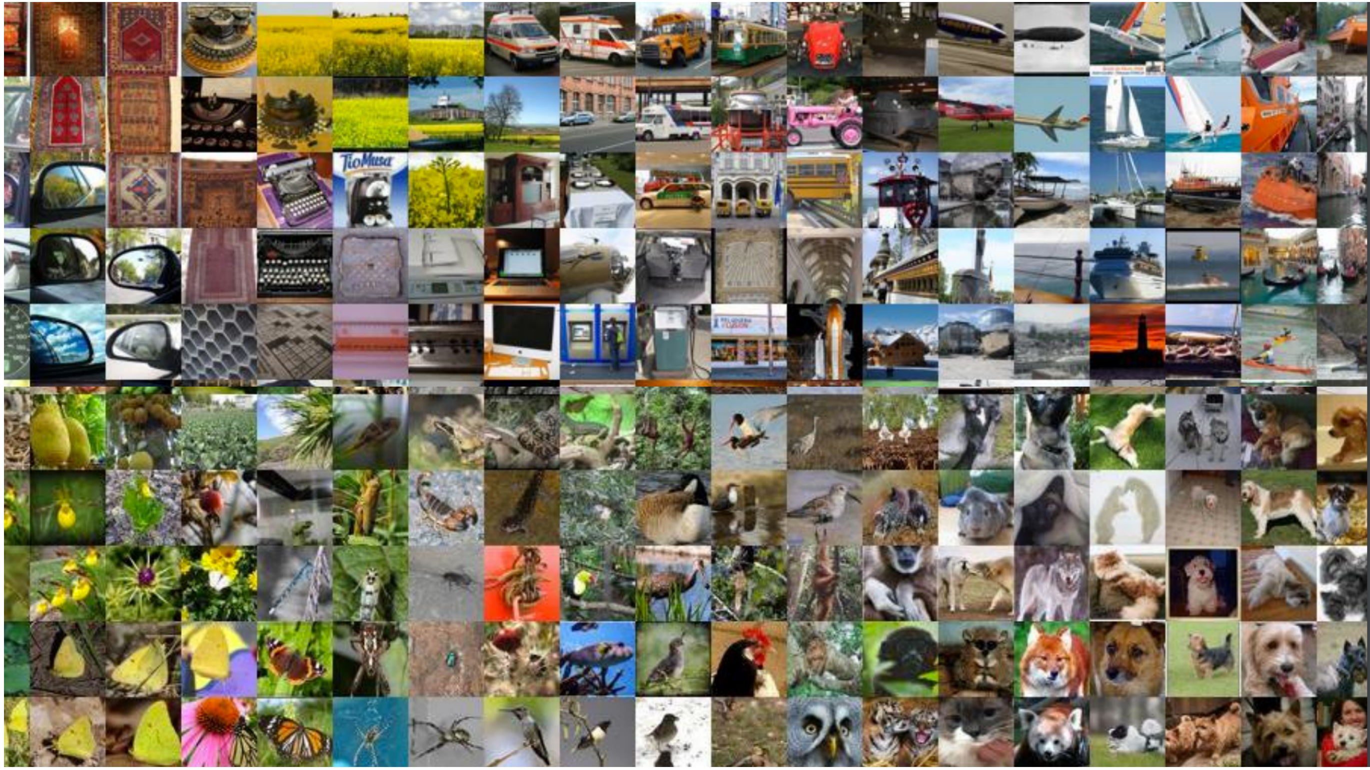
Microsoft

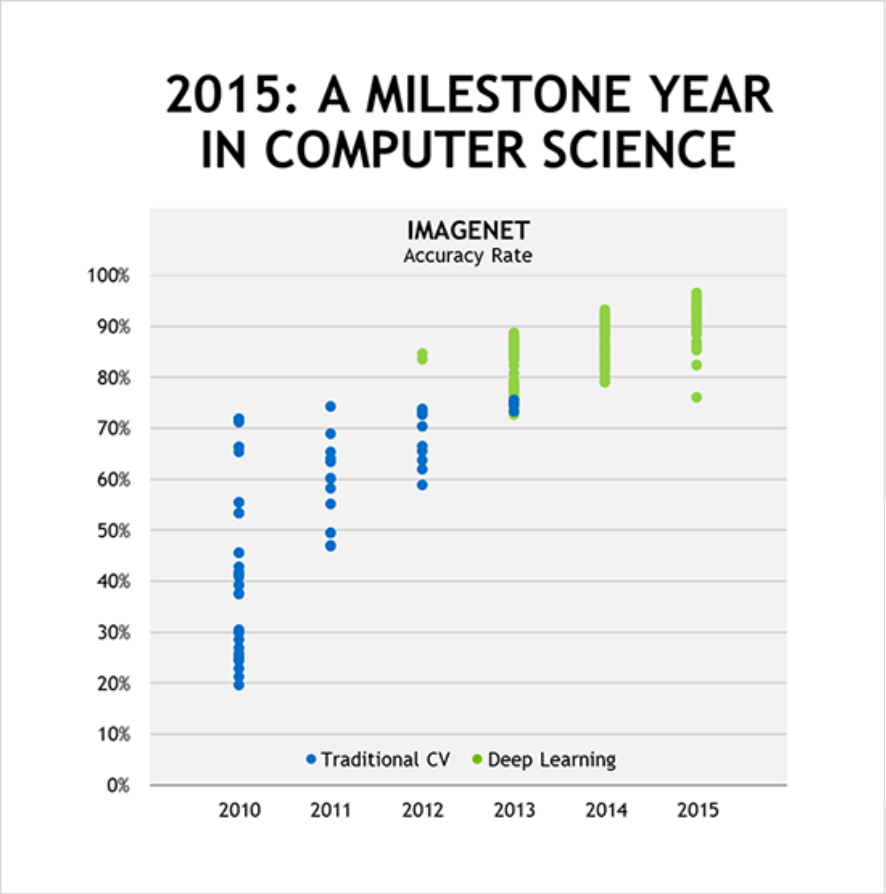
Images from EVIDIA

ImageNet

Released in 2009

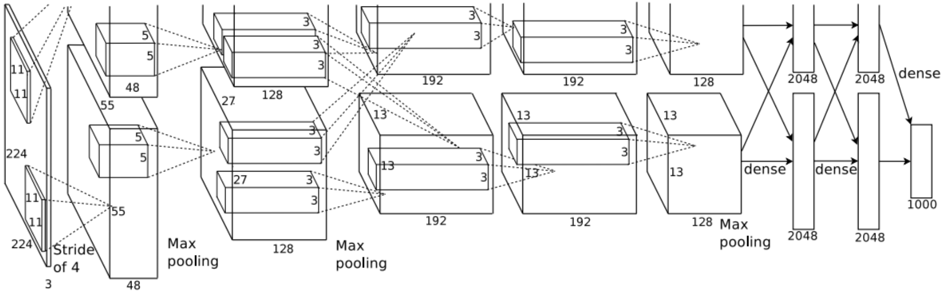
1.2 Million Images
More than 1000 classes





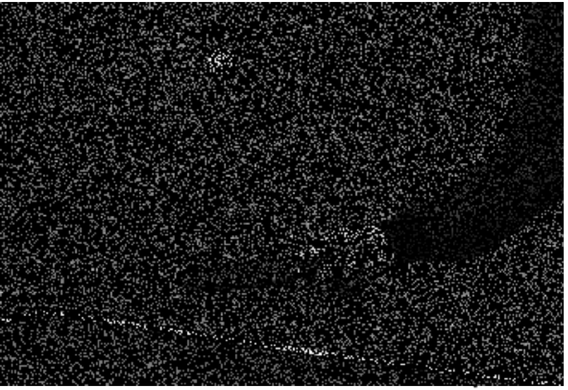
Alex Net

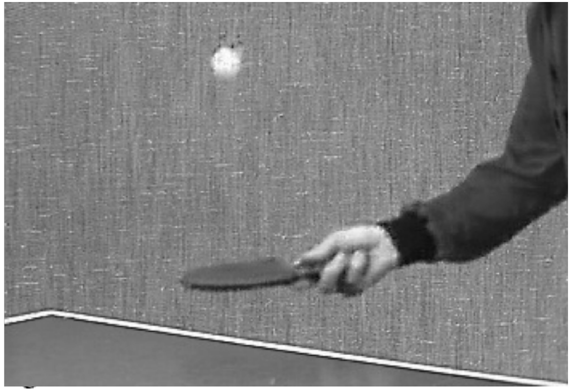
Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton 2012

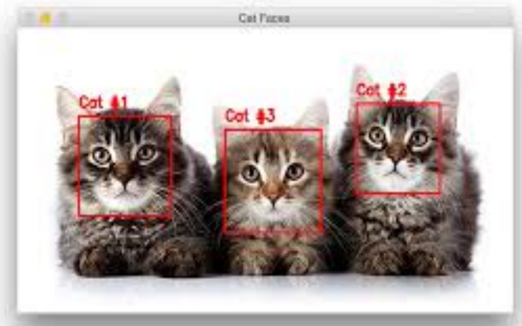


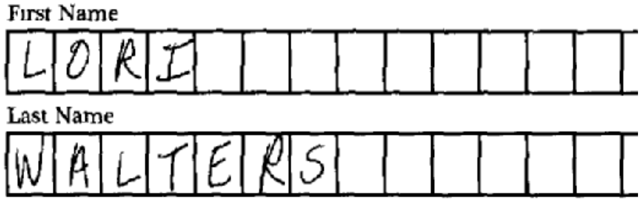
LeCunn: "Facebook uses networks with 50-100 layers"

Images from EVIDIA



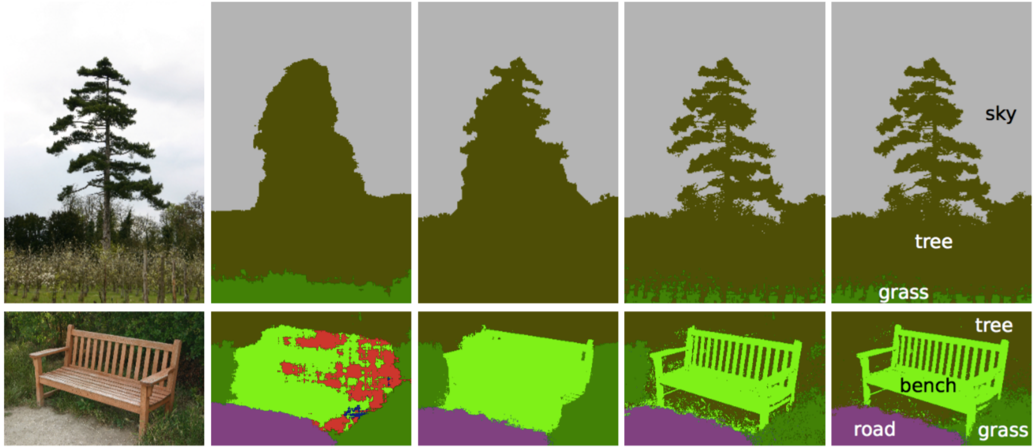






Images from EVIDIA DLI

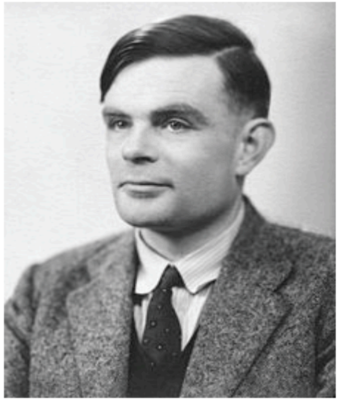
Semantic Segmentation

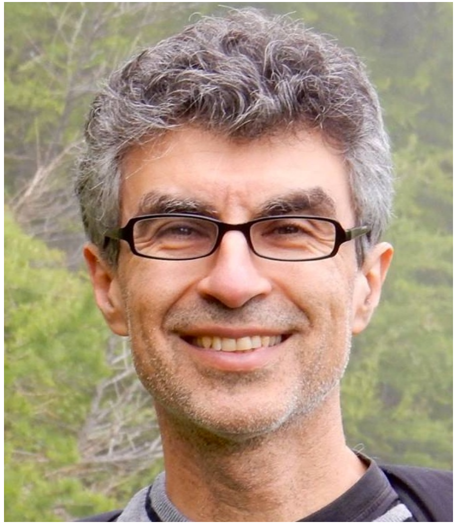


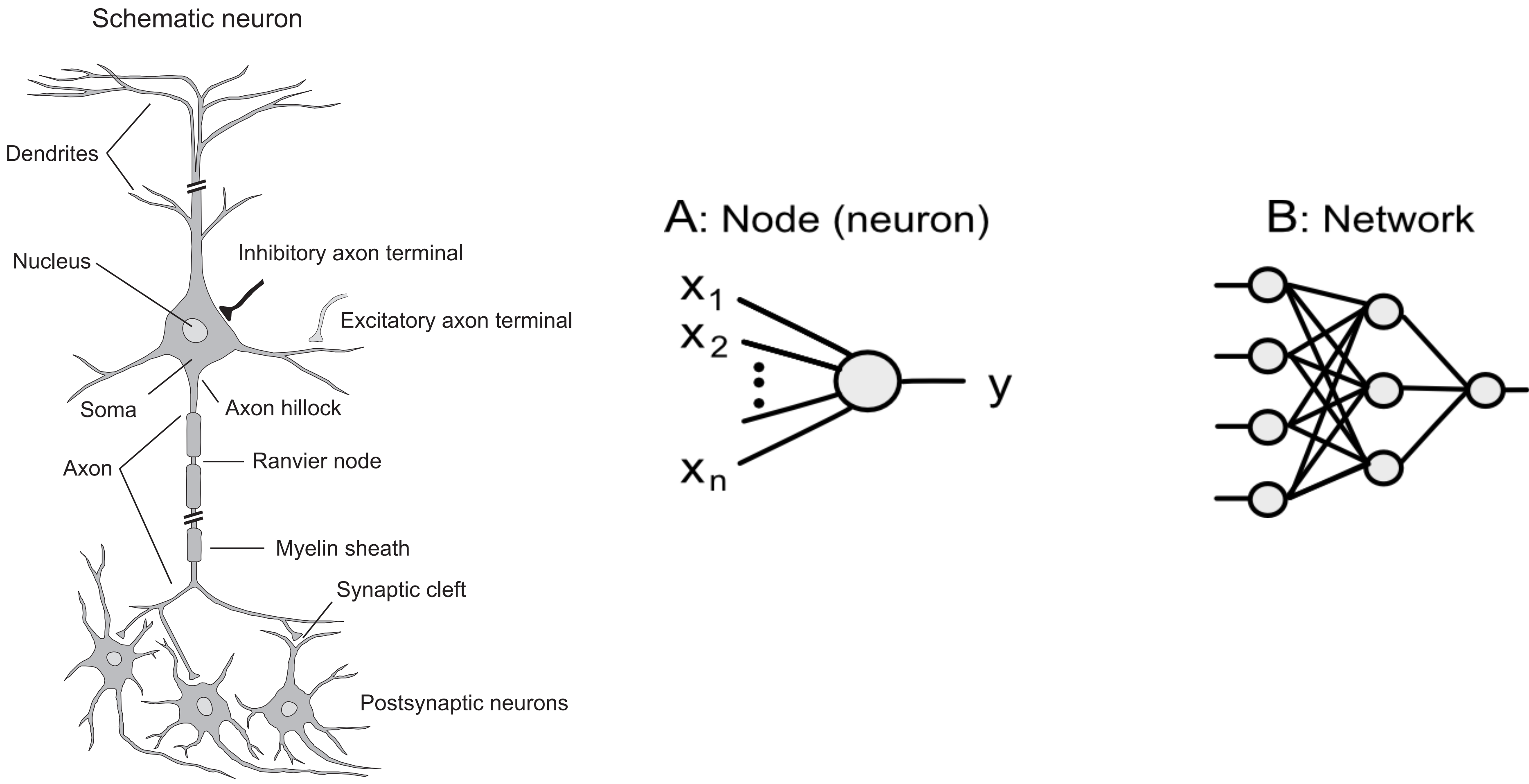
(a) Image (b) Unary classifiers (c) Robust P^n CRF (d) Fully connected CRF, MCMC inference, 36 hrs (e) Fully connected CRF, our approach, 0.2 seconds

http://blog.qure.ai/notes/semantic-segmentation-deep-learning-review

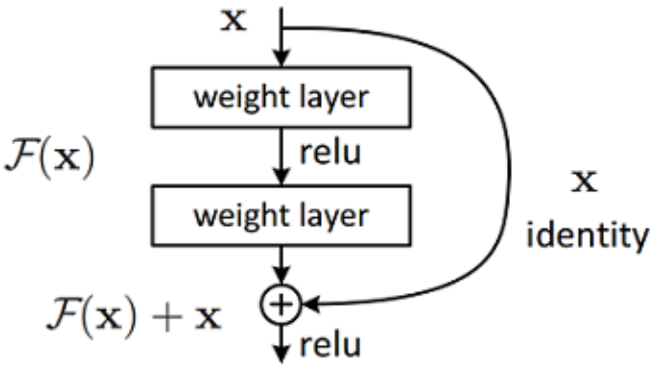
A little History



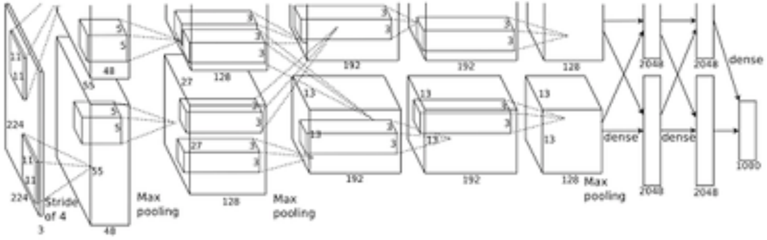



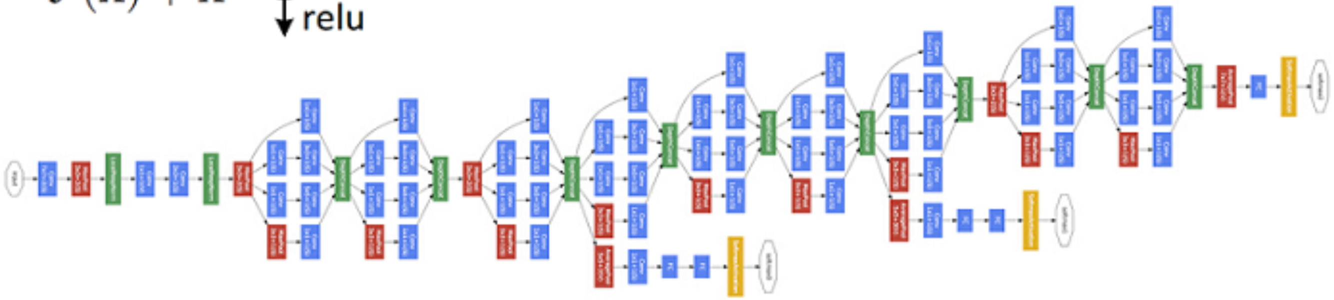
Schematic neuron
Dendrites
Nucleus
Inhibitory axon terminal
Excitatory axon terminal
Soma
Axon hillock
Axon
Ranvier node
Myelin sheath
Synaptic cleft
Postsynaptic neurons
A: Node (neuron)
x_1
x_2
x_n
y
B: Network

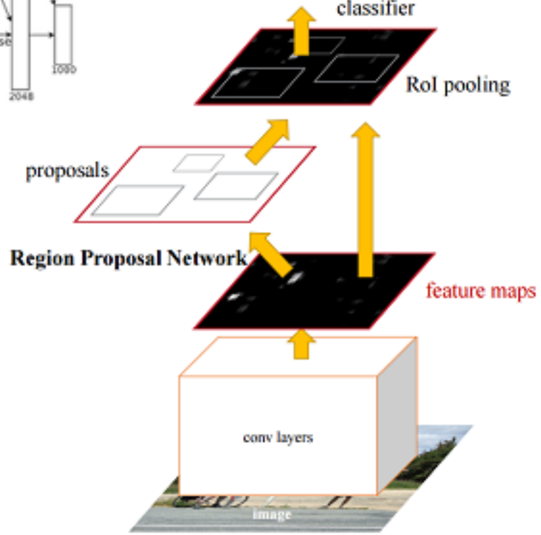


x
weight layer
relu
weight layer
$\mathcal{F}(\mathbf{x})$
x
identity
$\mathcal{F}(\mathbf{x}) + \mathbf{x}$
relu



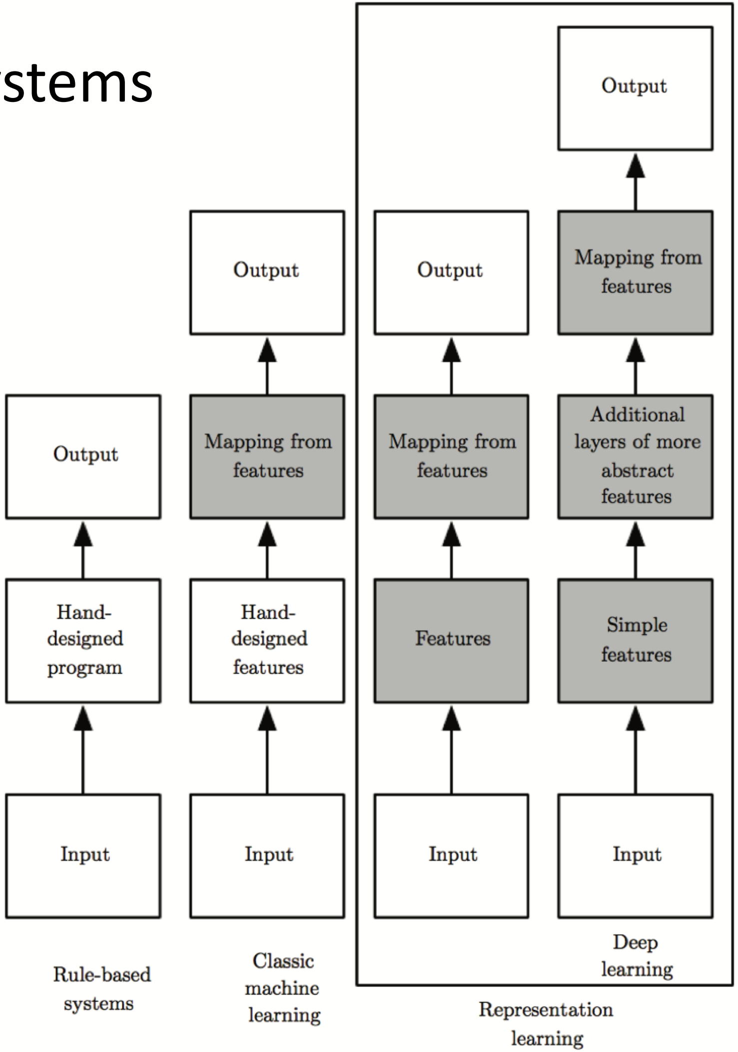
Stride of 4
Max pooling
Max pooling
Max pooling
dense
dense
dense





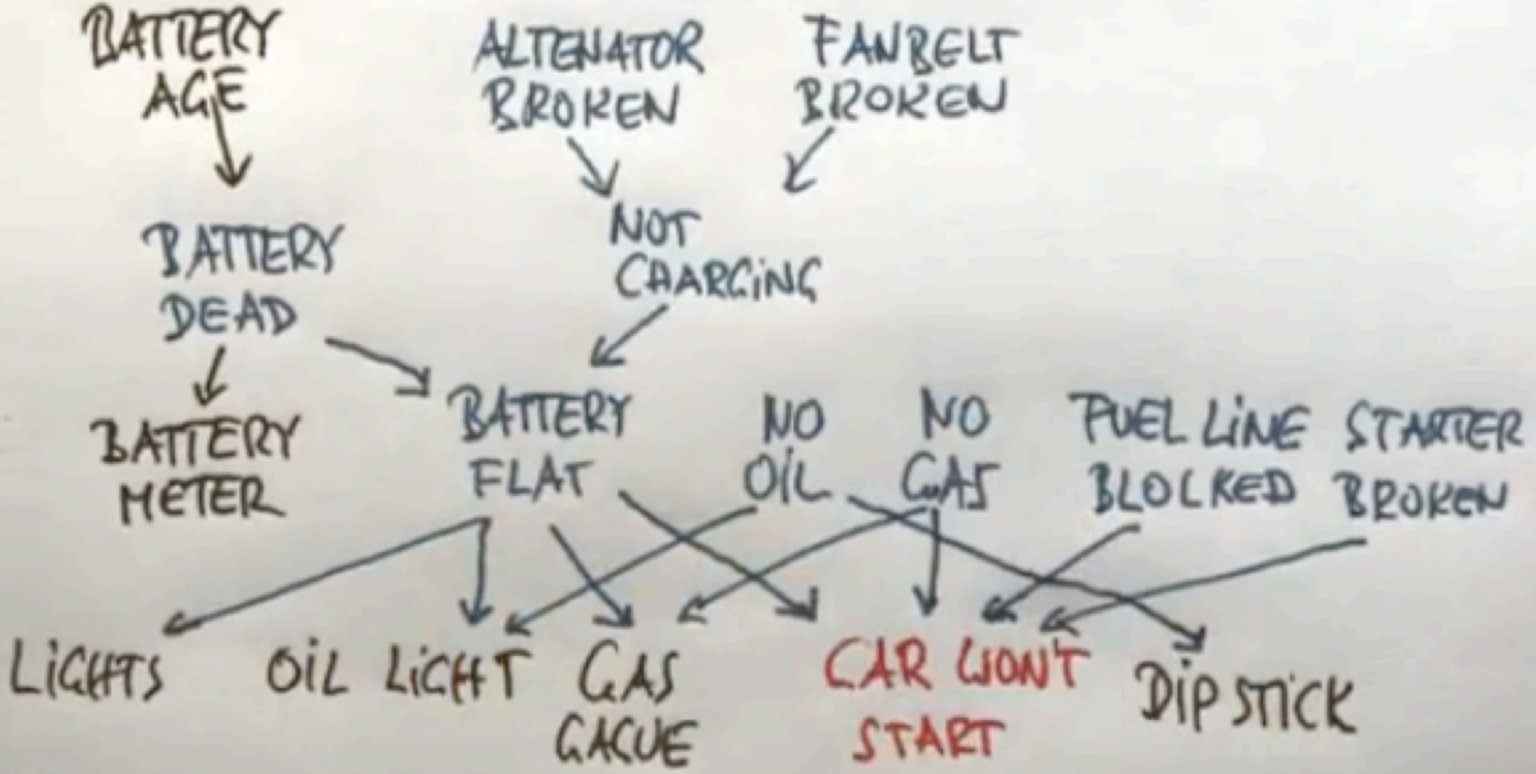
classifier
RoI pooling
proposals
Region Proposal Network
feature maps
conv layers
image

Evolution of machine learning systems



from Goodfellow, Bengio, Courville 2015

BAYES NETWORK



Symbolic vs subsymbolic AI

Symbolic programming
Explicit Rules
Ontologies

Neural Networks
Deep Learning
SVM, Decision Trees, LDA, …

Bayes

Supervised Learning

Unsupervised Learning

Reinforcement Leaning

Tool	Uses	Language
Scikit-Learn	Classification, Regression, Clustering	Python
Spark MLlib	Classification, Regression, Clustering	Scala, R, Java
Weka	Classification, Regression, Clustering	Java
Caffe	Neural Networks	C++, Python
TensorFlow	Neural Networks	Python

Mathematical representation

	Functional form	Probabilistic form
True underlying world:	$y = f(\mathbf{x})$	$p(Y = y\|\mathbf{x})$
Model:	$\hat{y} = \hat{f}(\mathbf{x}; \mathbf{w})$	$p(\hat{Y} = \hat{y}\|\mathbf{x}; \theta)$
Learning:	$w_i \leftarrow w_i - \alpha \bigtriangledown \mathcal{L}$	$\mathbf{w}^* = \text{argmax}_w p(\mathbf{w}\|y, \mathbf{x})$