Deep Learning (with TensorFlow 2.0)

An Interactive Introduction to Artificial Neural Networks

Jon Krohn, D.Phil.

#ODSC Open Data Science Conf. NY
June 28th, 2019

`github.com/jonkrohn/tf2`

POLL

What are you?

- Developer / Engineer
- Scientist / Analyst / Statistician / Mathematician
- Combination of the Above
- Other

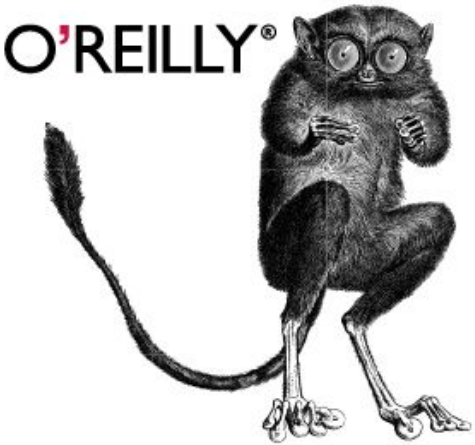
POLL

What's your level of experience with the topic?

- little to no exposure to deep learning
- some deep learning theory
- deep learning theory + experience with a deep learning library
- deep learning theory + experience with TensorFlow/Keras

untapt

Pearson

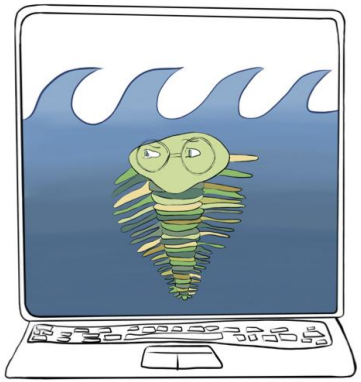


Pearson

ADDISON WESLEY DATA & ANALYTICS SERIES

DEEP LEARNING ILLUSTRATED

A Visual, Interactive Guide to Artificial Intelligence



JON KROHN

with **GRANT BEYLEVELD** *and* **AGLAÉ BASSENS**

Deep Learning Fundamentals

1. The Unreasonable Effectiveness of Deep Learning
2. Essential Deep Learning Theory
3. Deep Learning with TensorFlow 2.0

Deep Learning Fundamentals

Part 1:

The Unreasonable Effectiveness of Deep Learning

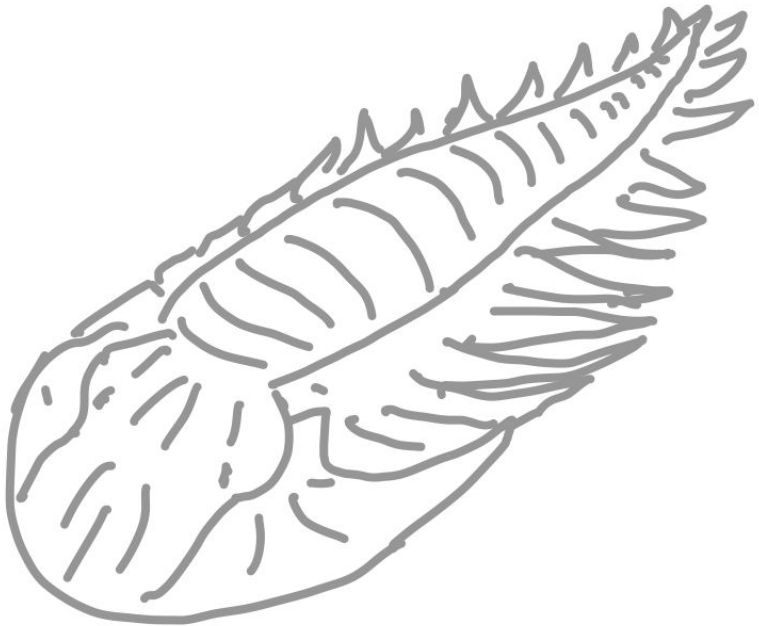
- Intro to Neural Networks and Deep Learning
- Deep Learning Families
- Deep Learning Libraries

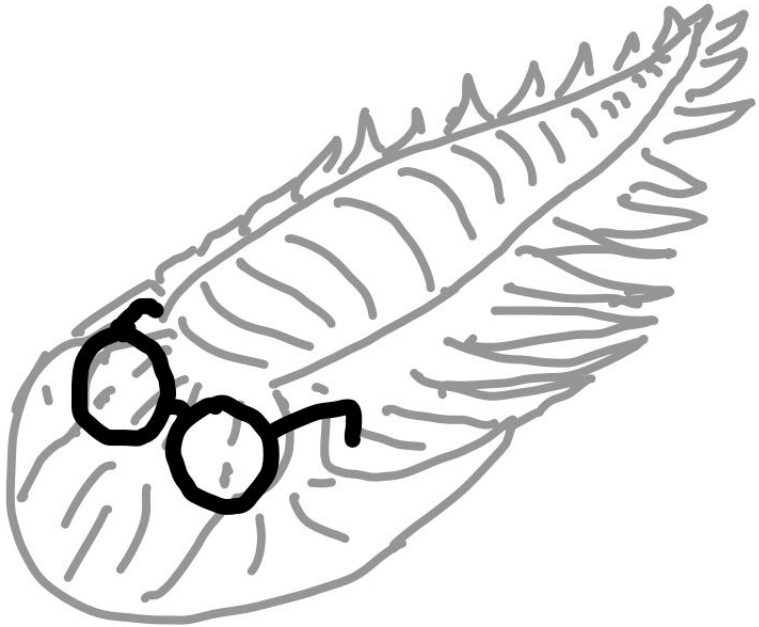
Deep Learning Fundamentals

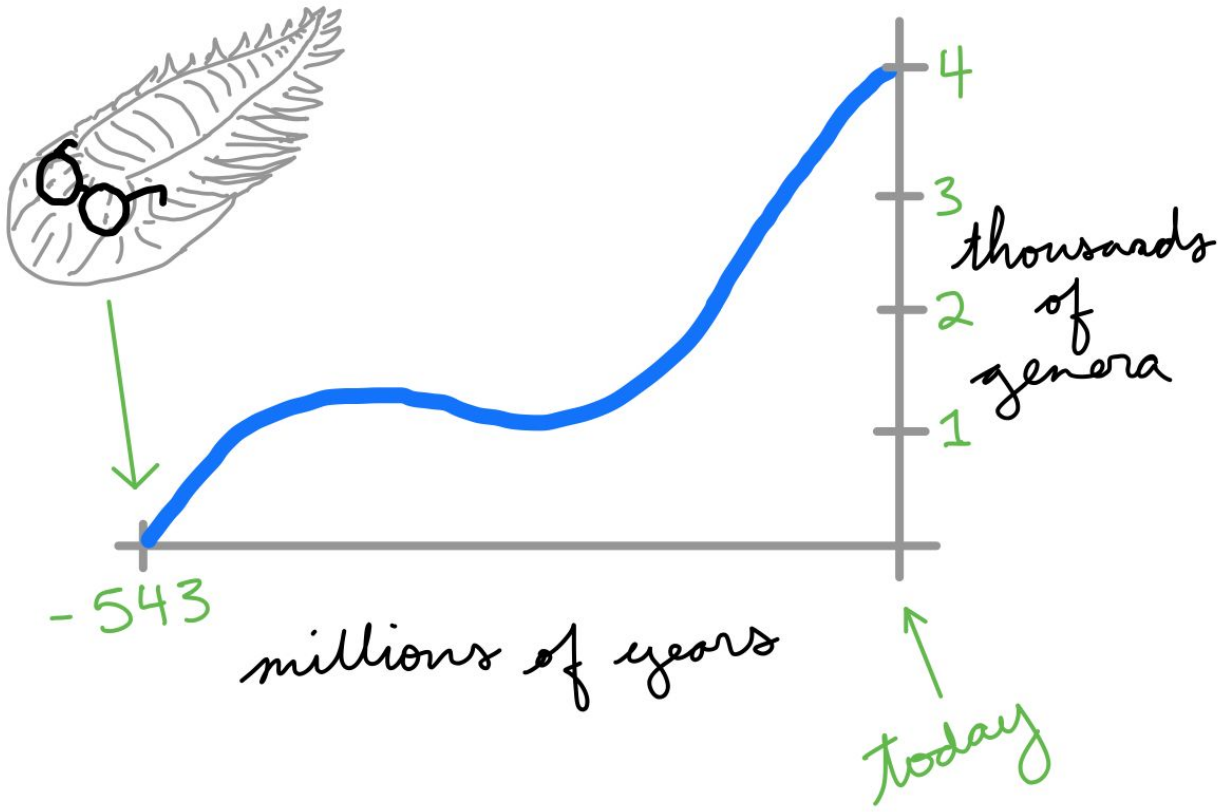
Part 1:

The Unreasonable Effectiveness of Deep Learning

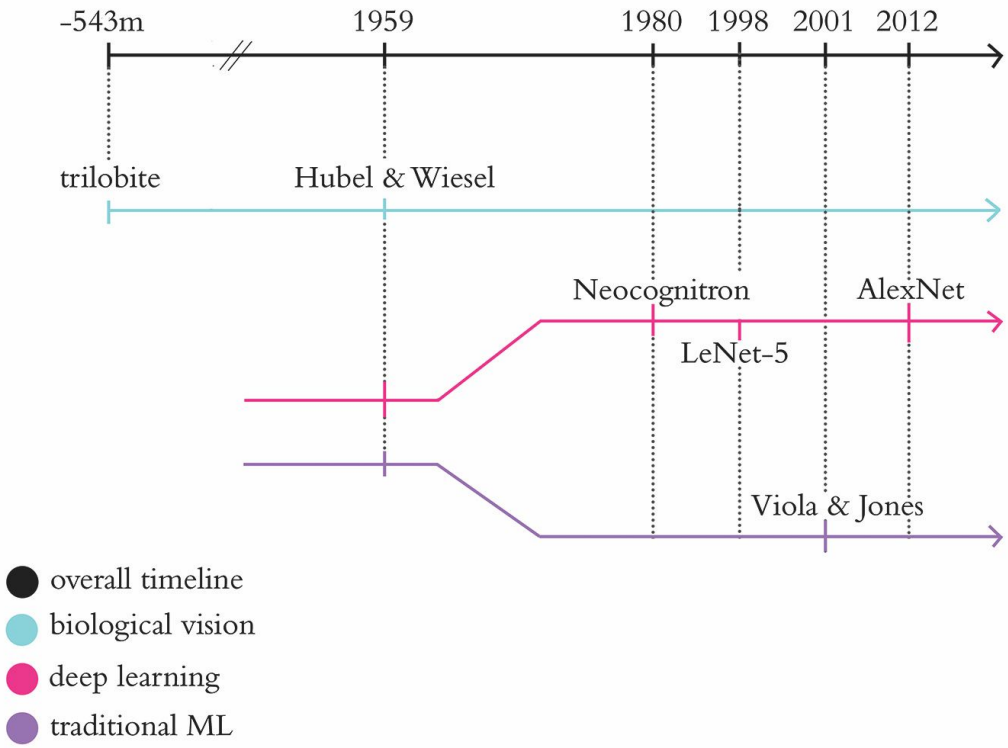
- **Intro to Neural Networks and Deep Learning** *(Chapter 1)*
- Deep Learning Families
- Deep Learning Libraries

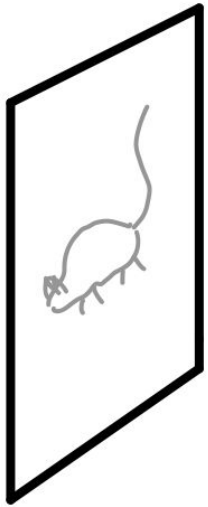


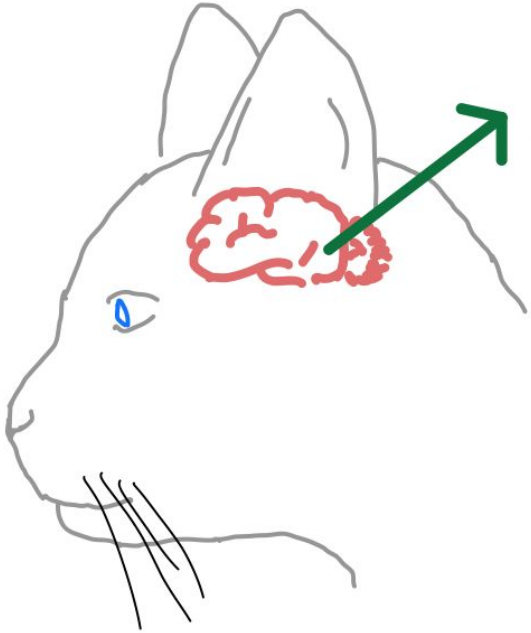


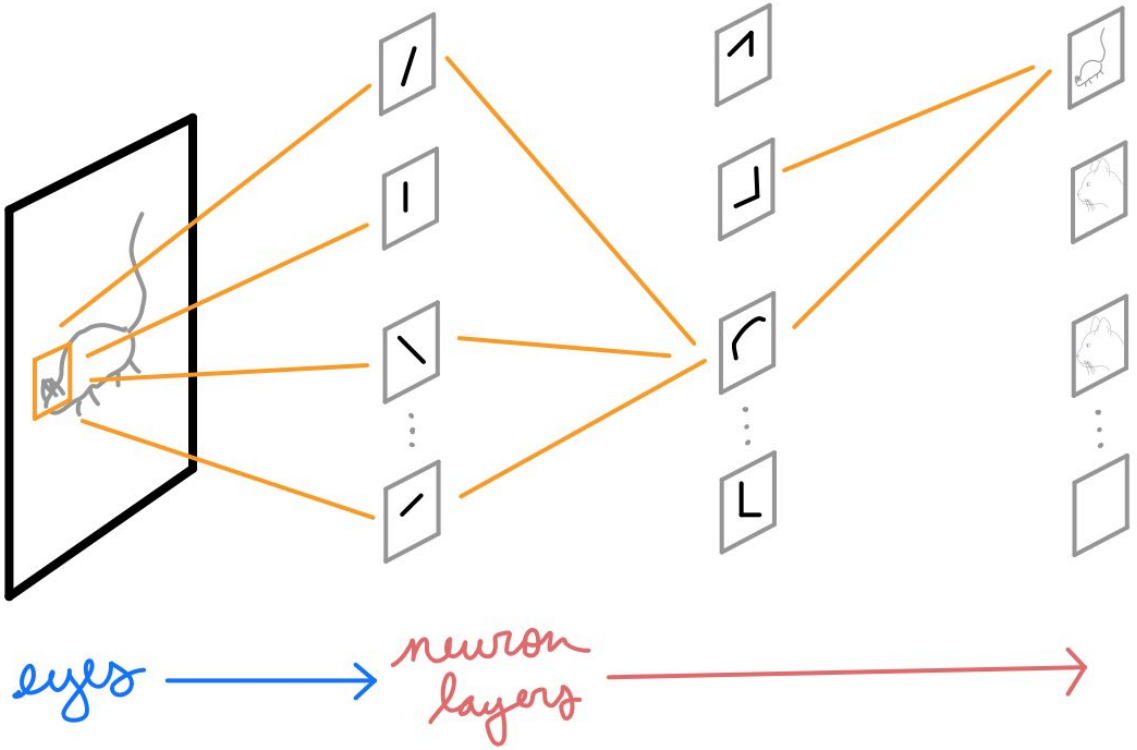


Case Study: The History of Vision

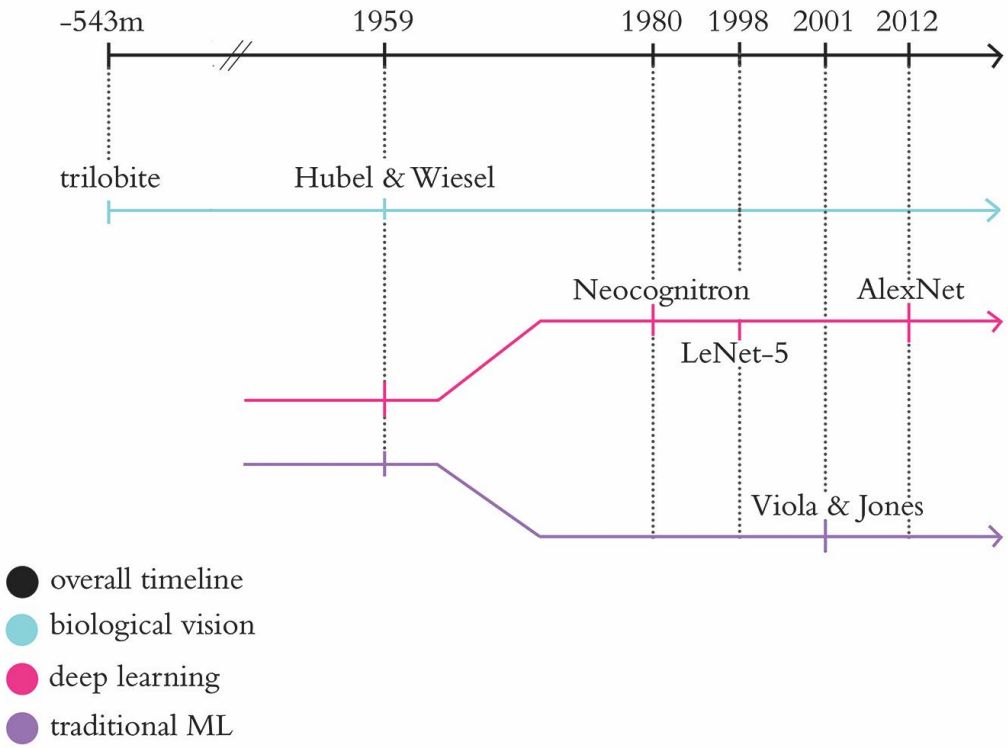








Case Study: The History of Vision



Neocognitron (Fukushima, 1980)

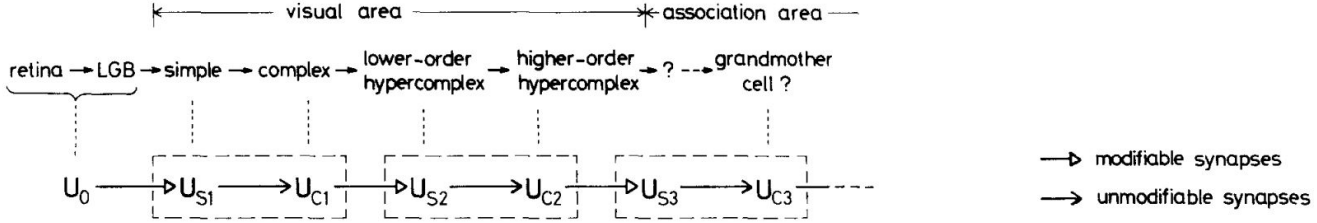


Fig. 1. Correspondence between the hierarchy model by Hubel and Wiesel, and the neural network of the neocognitron

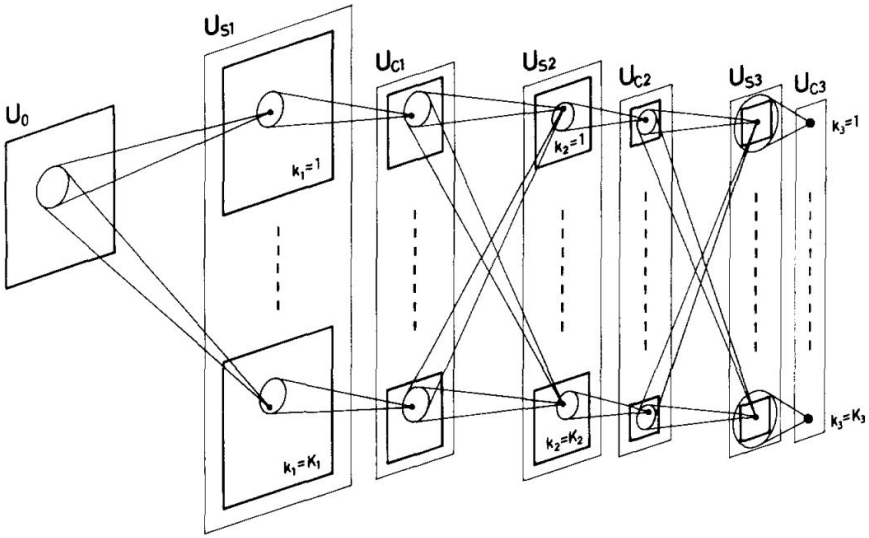
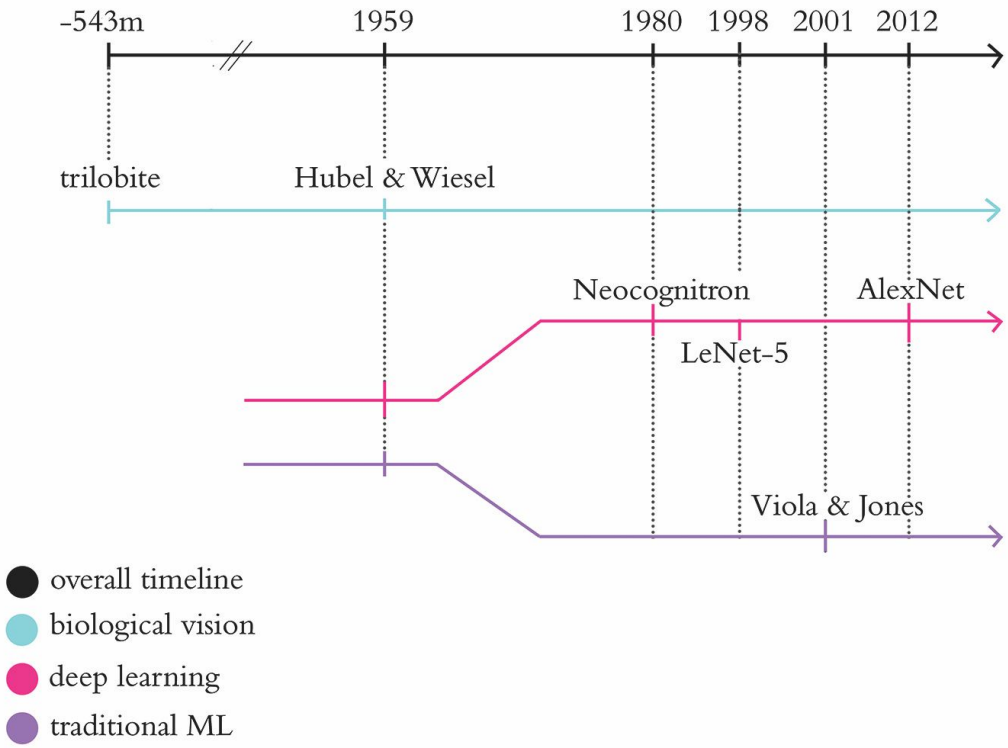
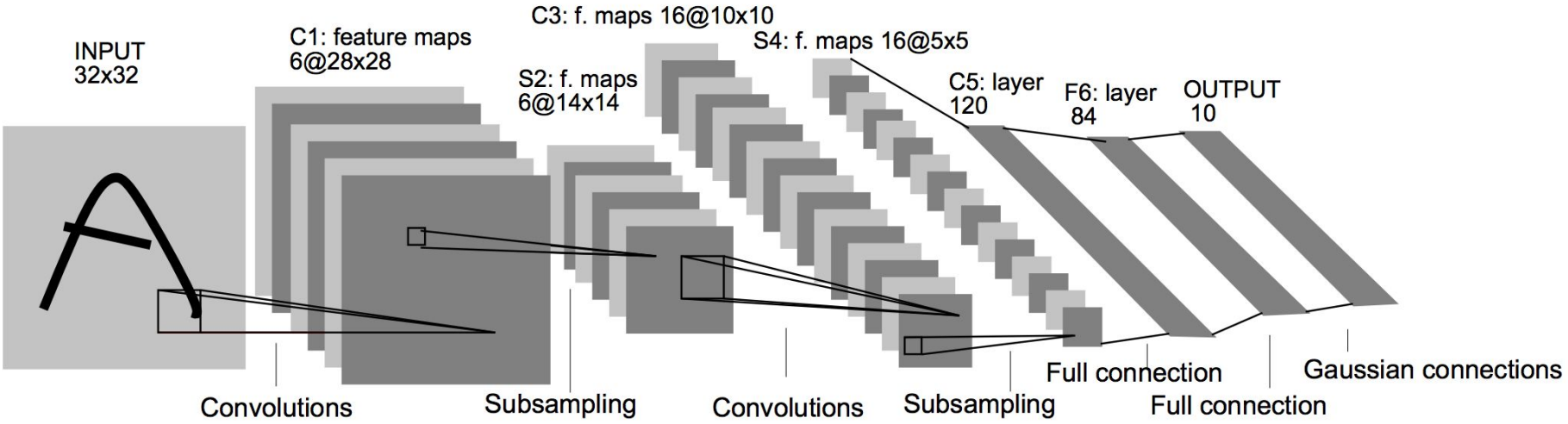


Fig. 2. Schematic diagram illustrating the interconnections between layers in the neocognitron

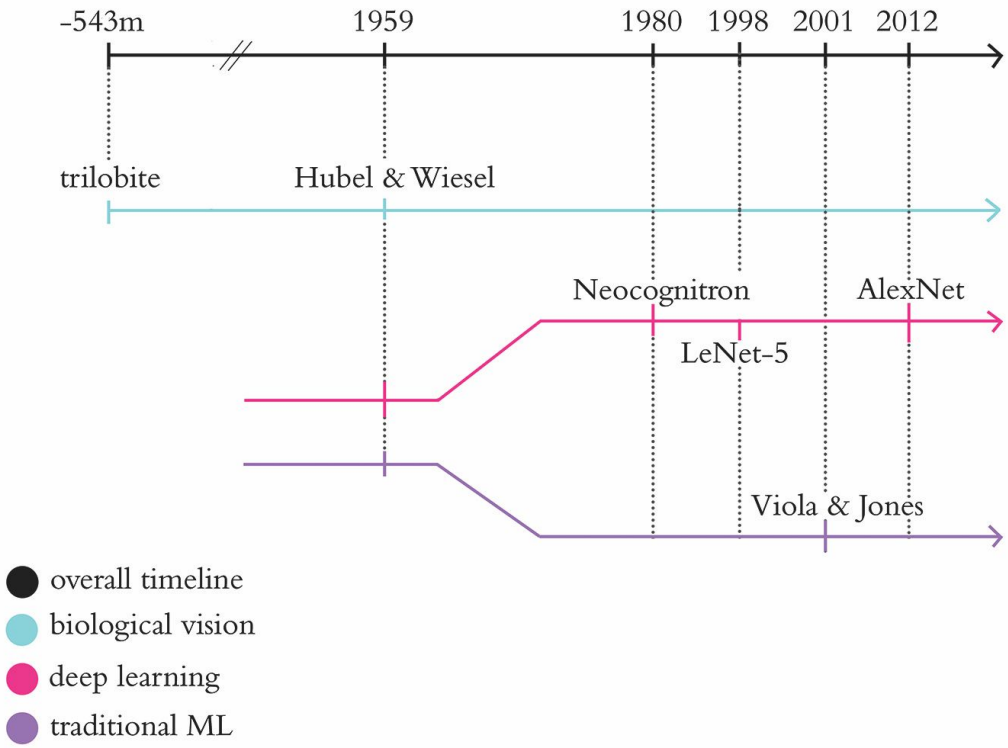
Case Study: The History of Vision



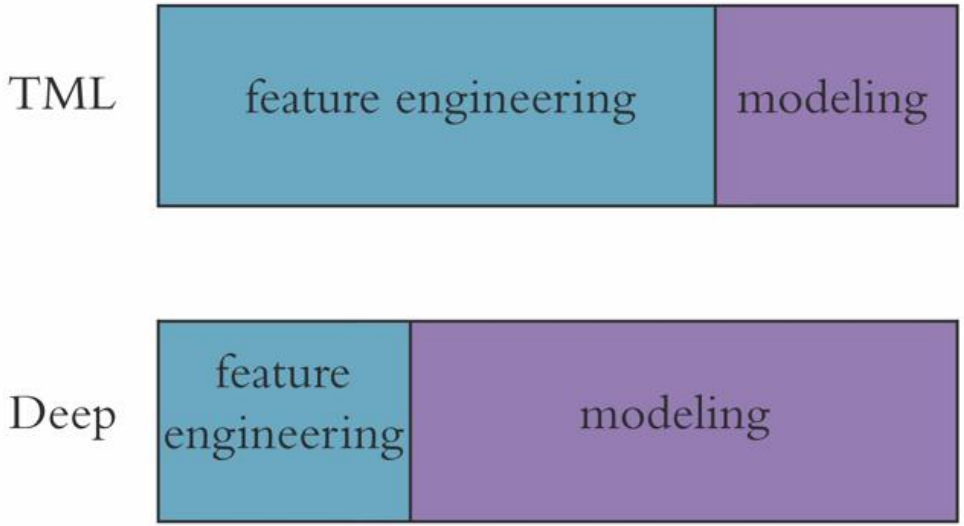
LeNet-5 (LeCun et al., 1998)



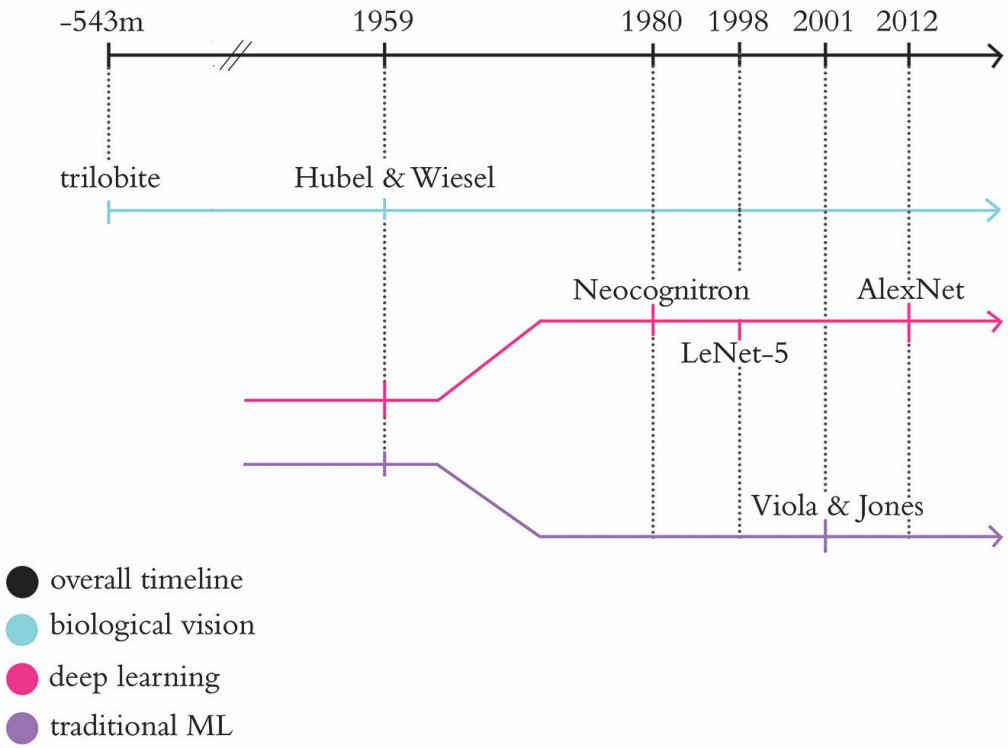
Case Study: The History of Vision



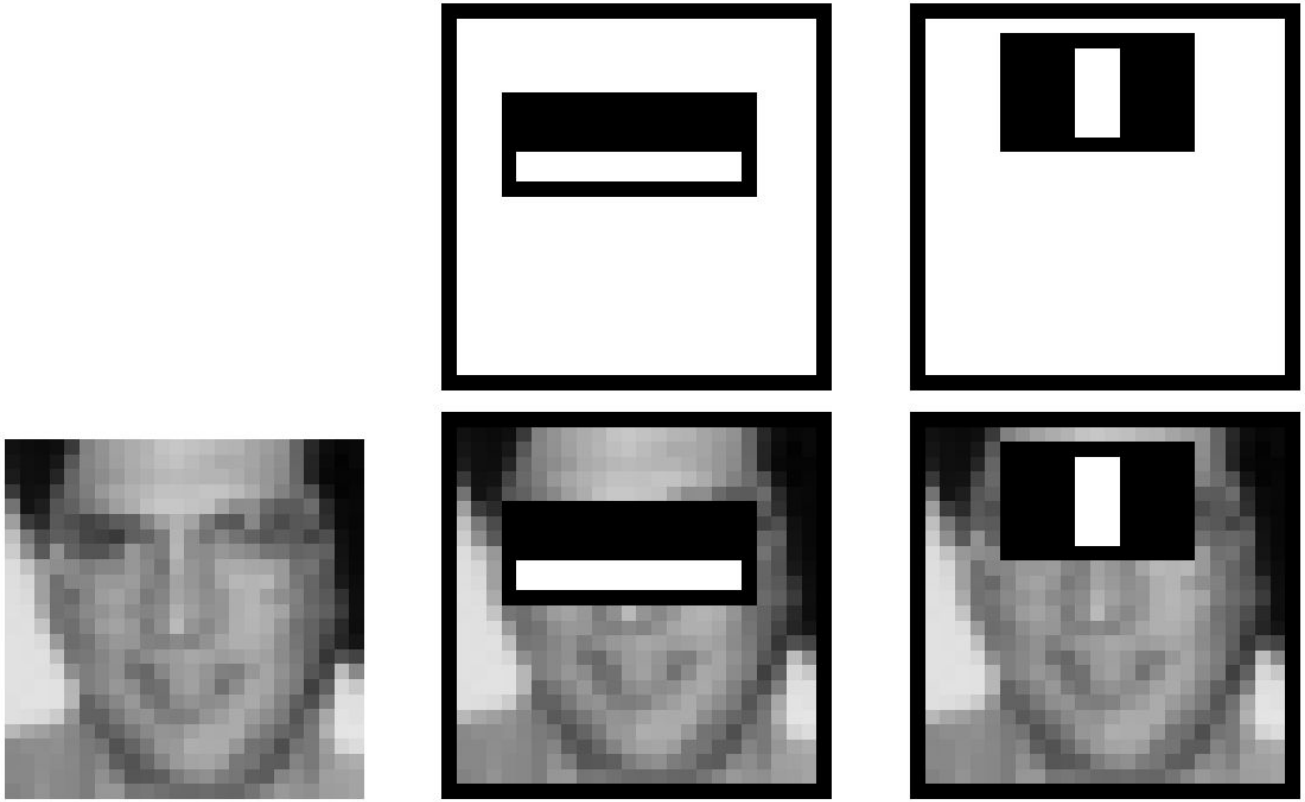
Traditional ML vs Deep Learning



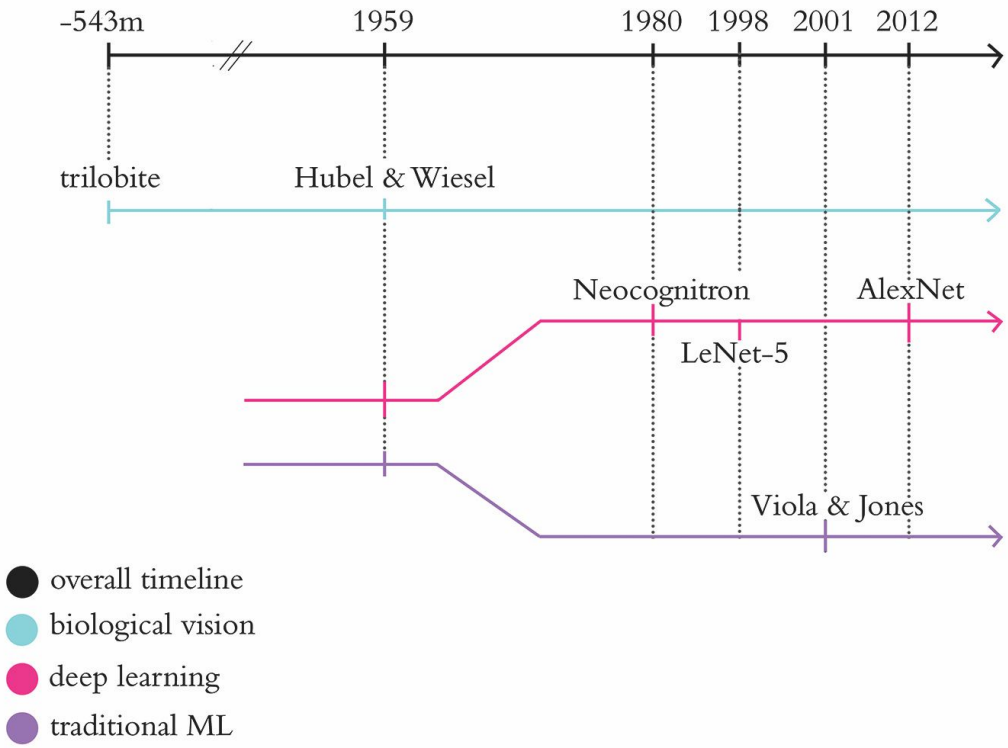
Case Study: The History of Vision



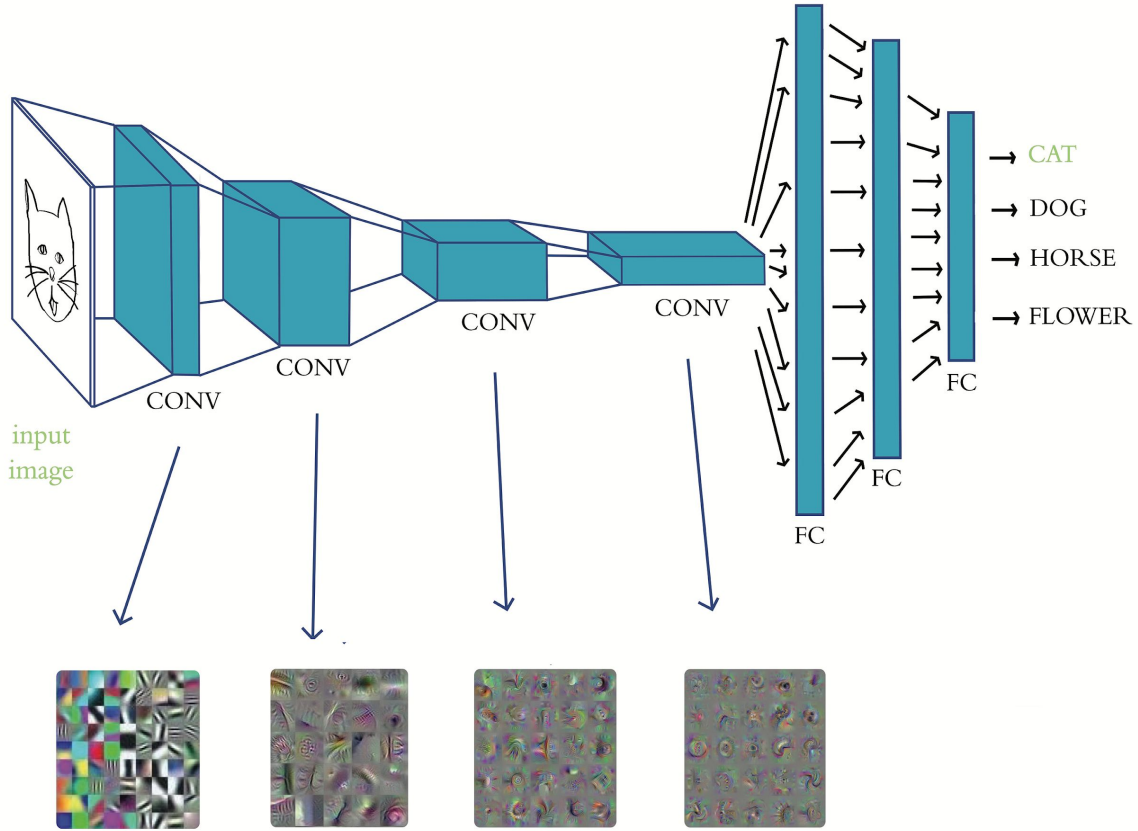
Viola & Jones (2001)



Case Study: The History of Vision



AlexNet (Krizhevsky et al., 2012)



POLL

If a voice recognition algorithm is fed audio of speech as inputs, given corresponding text as the outputs (labels) to learn, and no features are explicitly programmed, is this a:

- Traditional Machine Learning Algorithm
- Deep Learning Algorithm
- I Don't Know

Deep Learning Fundamentals

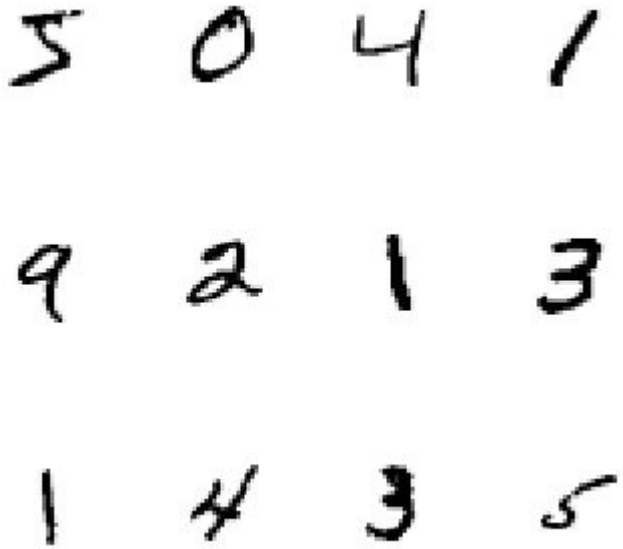
Part 1:

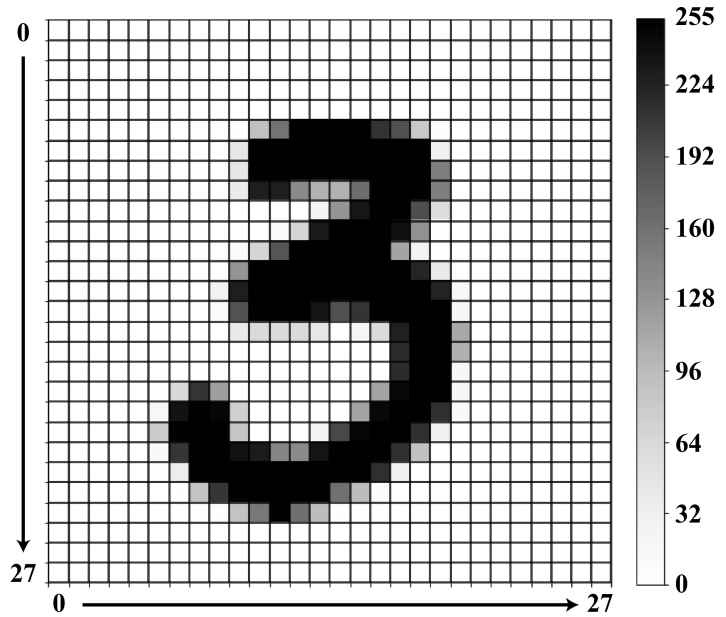
The Unreasonable Effectiveness of Deep Learning

- Intro to Neural Networks and Deep Learning
- **Deep Learning Families** *(Chapters 2-4)*
- Deep Learning Libraries

Dense Networks

The Cart Before the Horse *(Chapter 5)*



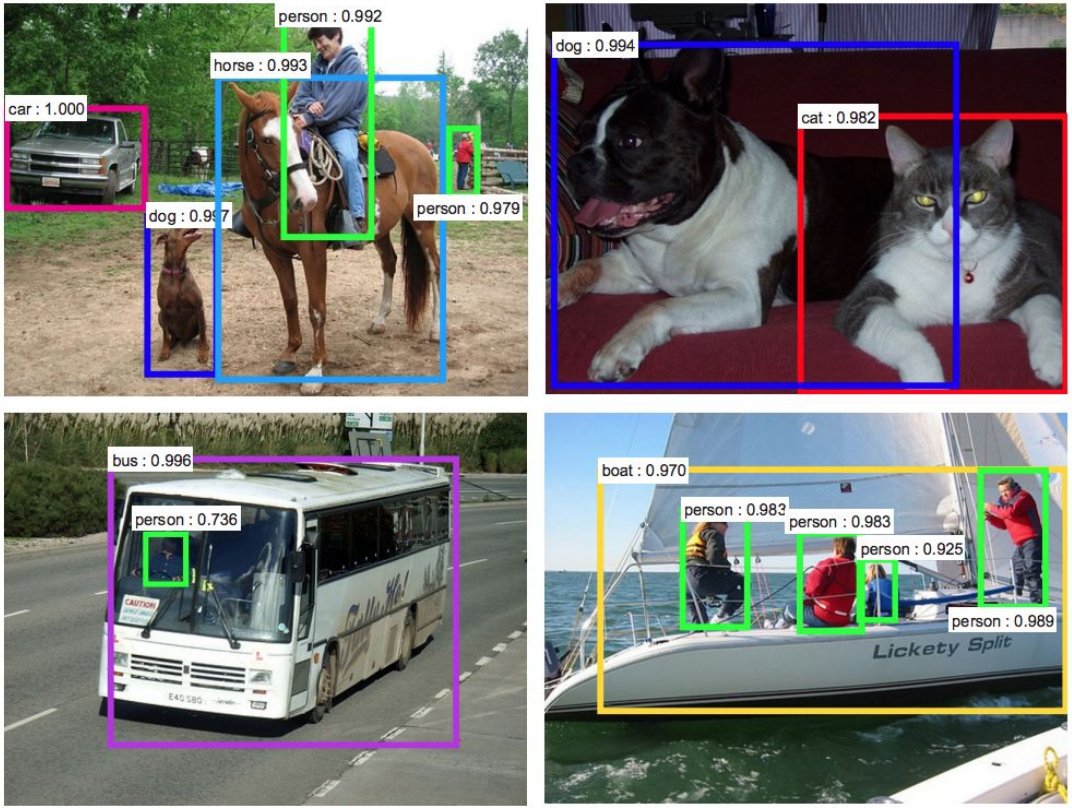


interactive Colab demo: *Shallow Net in TF 2.0* (`bit.ly/shallowTF`)

GitHub repo: `github.com/jonkrohn/tf2`

ConvNets: Convolutional Networks

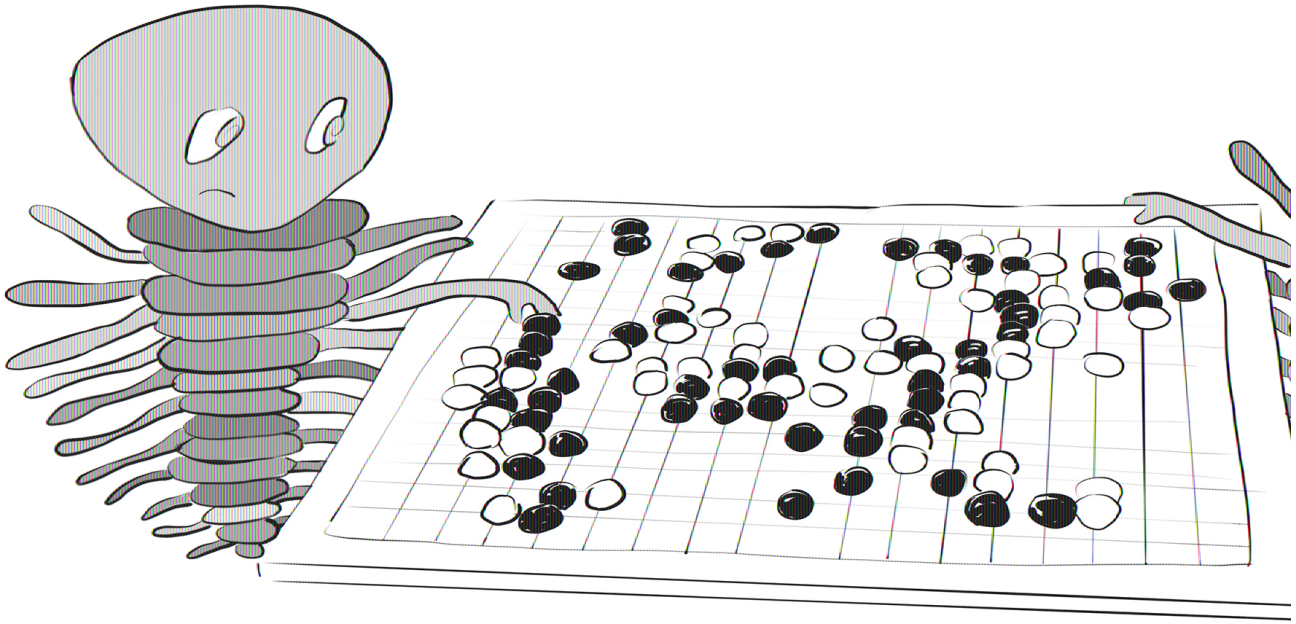
ConvNets: Convolutional Networks



Ren et al. (2015)

RNNs: Recurrent Neural Networks

Deep Reinforcement Learning



GANs: Generative Adversarial Networks

GANs: Generative Adversarial Networks

Karros et al. (2018)

POLL

If you were designing an algorithm to learn to play Tetris by maximizing its score, which of these Deep Learning approaches would be most appropriate?

- Convolutional Neural Network
- Recurrent Neural Network
- Deep Reinforcement Learning
- Generative Adversarial Network

POLL

If you were designing an algorithm to recognise tumours in medical images, which of these Deep Learning approaches would be most appropriate?

- Convolutional Neural Network
- Recurrent Neural Network
- Deep Reinforcement Learning
- Generative Adversarial Network

POLL

If you were designing an algorithm to predict stock price movements based on time series data, which of these Deep Learning approaches would be most appropriate?

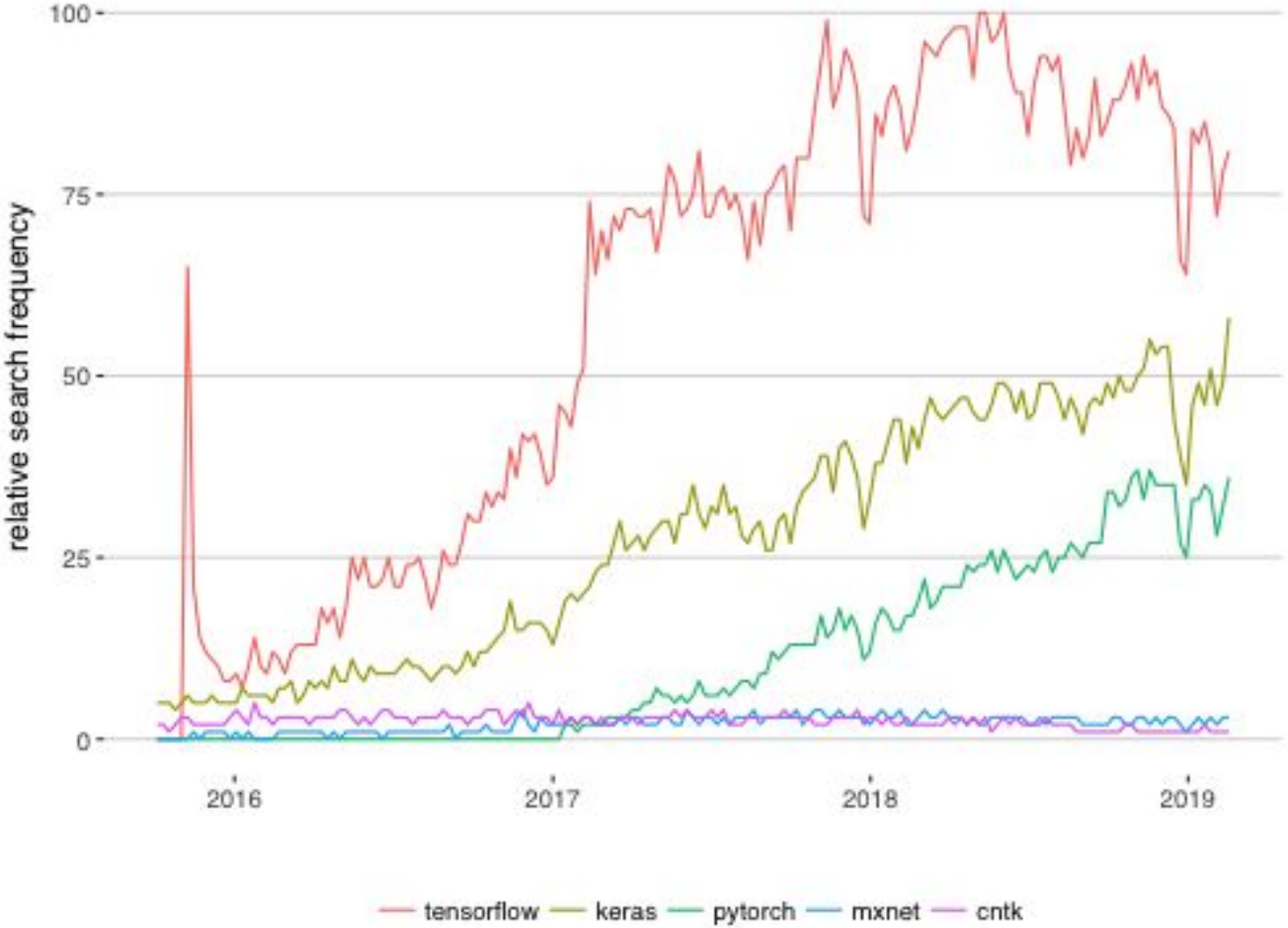
- Convolutional Neural Network
- Recurrent Neural Network
- Deep Reinforcement Learning
- Generative Adversarial Network

Deep Learning Fundamentals

Part 1:

The Unreasonable Effectiveness of Deep Learning

- Intro to Neural Networks and Deep Learning
- Deep Learning Families
- **Deep Learning Libraries** *(Chapter 14)*



Pearson

Leading Deep Learning Libraries

	Caffe	Torch	MXNet	TensorFlow
Language	Python, Matlab	Lua, C	Python, R, C++ Julia, Matlab JavaScript, Go Scala, Perl	**Python, C**, C++ Java, Go, JS, Swift (*Haskell, Julia, R, Scala, Rust, C#)*
Programming Style	Symbolic	Imperative	Imperative	Symbolic… *for now*
Parallel GPUs: Data	Yes	Yes	Yes	Yes
Parallel GPUs: Model		Yes	Yes	Yes
Pre-Trained Models	Model Zoo	Model Zoo	Model Zoo	github.com/tensorflow/models
High-Level APIs		PyTorch	in-built	Keras
Particular Strength	CNNs	interactivity		production deployment

Deep Learning Fundamentals

Part 2:

Essential Deep Learning Theory

- **Learning with Artificial Neurons** *(Chapters 6-7)*
- **TensorFlow Playground**

“Whiteboarding”!

Your Arsenal

Neurons

- sigmoid
- tanh
- ReLU

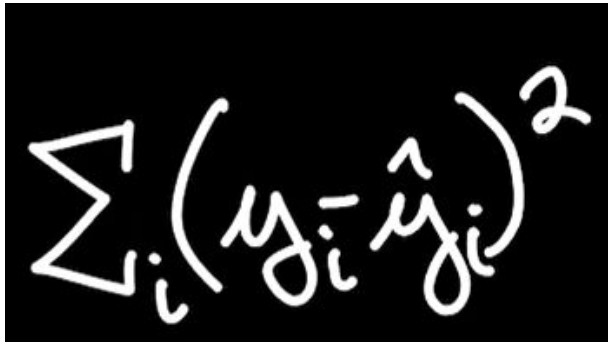
Your Arsenal

Neurons

- sigmoid
- tanh
- ReLU

Cost Functions

- quadratic cost



Your Arsenal

Neurons

- sigmoid
- tanh
- ReLU

Cost Functions

- quadratic cost
- cross-entropy

Your Arsenal

Neurons

- sigmoid
- tanh
- ReLU

Cost Functions

- quadratic cost
- cross-entropy

Gradient Descent

Your Arsenal

Neurons

- sigmoid
- tanh
- ReLU

Cost Functions

- quadratic cost
- cross-entropy

Gradient Descent

Backpropagation

Your Arsenal

Neurons

- sigmoid
- tanh
- ReLU

Cost Functions

- quadratic cost
- cross-entropy

Gradient Descent

Backpropagation

Layers

- dense
- softmax

Your Arsenal

Neurons

- sigmoid
- tanh
- ReLU

Cost Functions

- quadratic cost
- cross-entropy

Gradient Descent

Backpropagation

Layers

- dense
- softmax

Initialization

- Glorot

Your Arsenal

Neurons

- sigmoid
- tanh
- ReLU

Cost Functions

- quadratic cost
- cross-entropy

***Stochastic* Gradient Descent**

- mini-batch size
- learning rate
- second-order, e.g., Adam

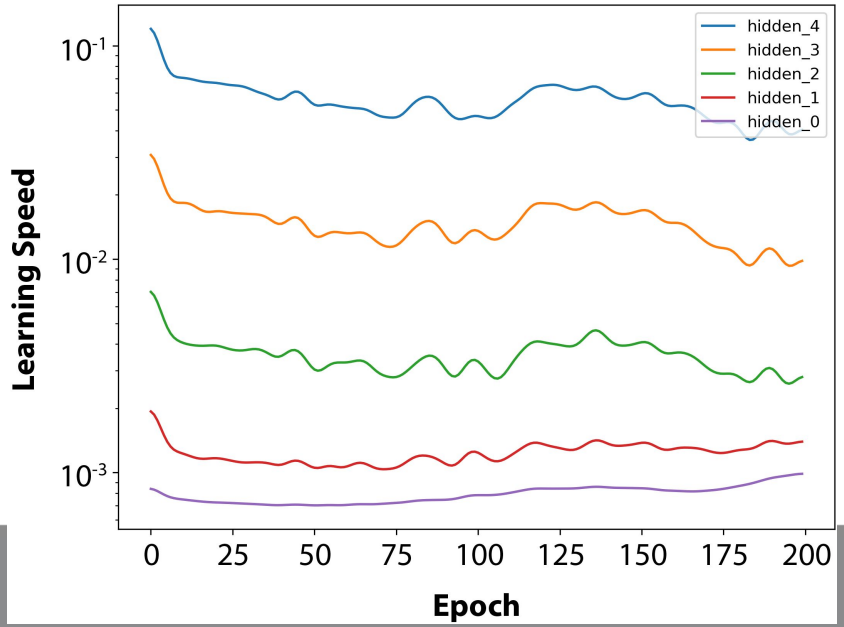
Backpropagation

Initialization

- Glorot

Layers

- dense
- softmax



Your Arsenal

Neurons

- sigmoid
- tanh
- ReLU

Cost Functions

- quadratic cost
- cross-entropy

***Stochastic* Gradient Descent**

- mini-batch size
- learning rate
- second-order, e.g., Adam

Backpropagation

Initialization

- Glorot

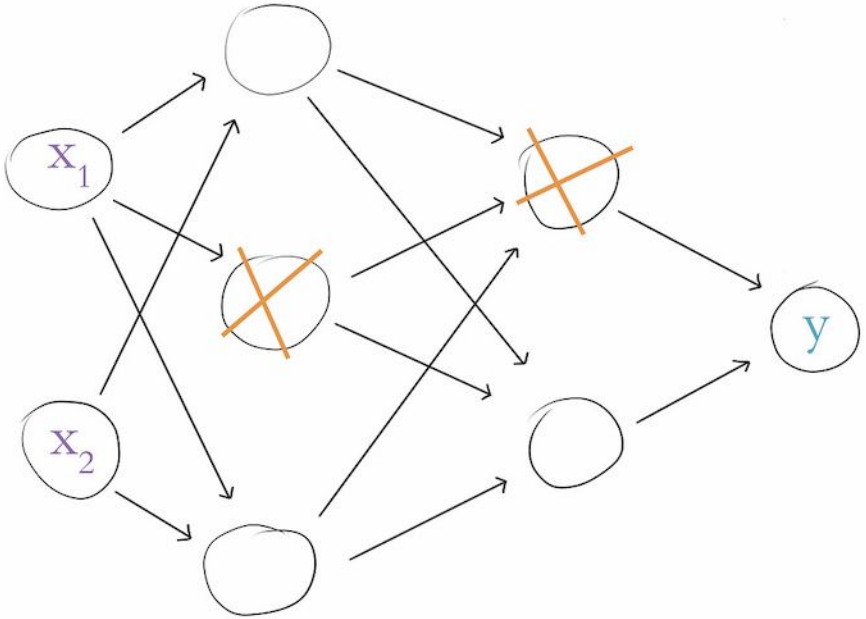
Layers

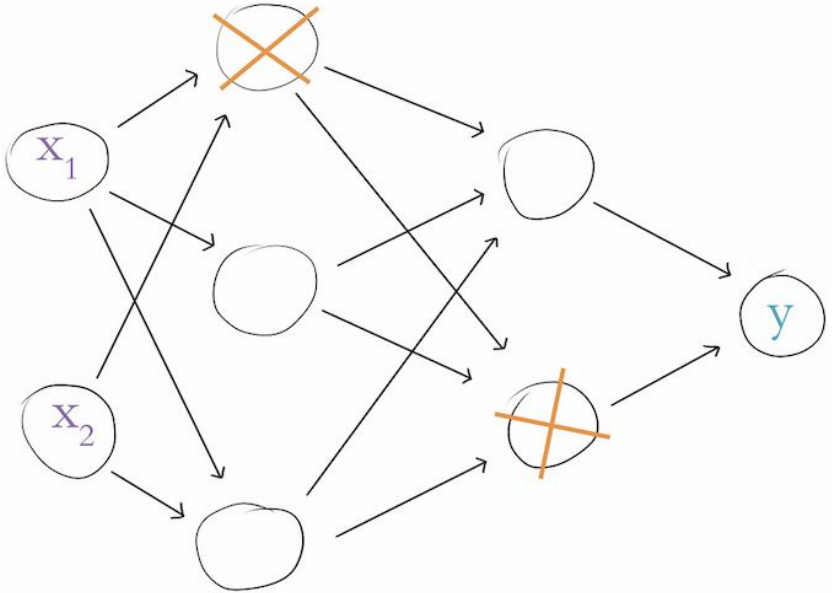
- dense
- softmax

Avoiding Overfitting

- Dropout

Dropout





Your Arsenal

Neurons

- sigmoid
- tanh
- ReLU

Cost Functions

- quadratic cost
- cross-entropy

***Stochastic* Gradient Descent**

- mini-batch size
- learning rate
- second-order, e.g., Adam

Backpropagation

Initialization

- Glorot

Layers

- dense
- softmax

Avoiding Overfitting

- Dropout
- Data Augmentation

TensorFlow Playground

interactive demo: *playground.tensorflow.org*

Deep Learning Fundamentals

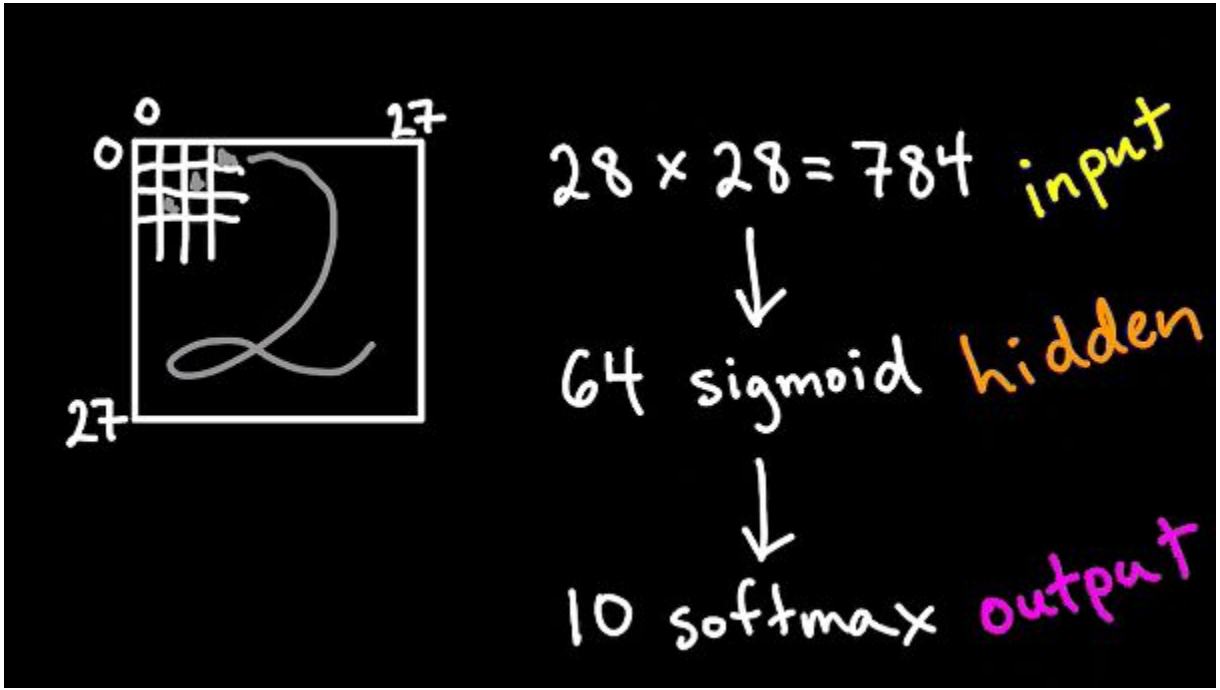
Part 3:

Deep Learning with TensorFlow

- **Revisiting our Shallow Neural Network**
- **Deep Nets in TensorFlow** *(Chapters 8-9)*
- **What to Study Next, Depending on Your Interests**

Revisiting our Shallow Net

interactive Colab demo: *Shallow Net in TensorFlow*



Deep Nets in TensorFlow

interactive Colab demo: *Deep Net in TF 2.0* (`bit.ly/deepNetTF`)

POLL

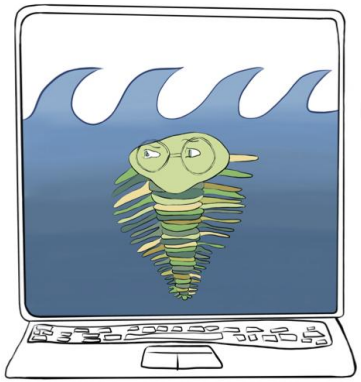
What follow-up Deep Learning topics interest you most?

- CNNs and Machine Vision
- Sequences: RNNs, LSTMs, NLP, Financial Time Series
- Generative Adversarial Networks
- Deep Reinforcement Learning
- Something Else

ADDISON WESLEY DATA & ANALYTICS SERIES

DEEP LEARNING ILLUSTRATED

A Visual, Interactive Guide to Artificial Intelligence



JON KROHN

with **GRANT BEYLEVELD** *and* **AGLAÉ BASSENS**

son

Stay in Touch

twitter.com/JonKrohnLearns

medium.com/@jonkrohn

linkedin.com/in/jonkrohn

(with message mentioning ODSC)