Machine Learning
«A gentle Introduction»

@ZimMatthias

Matthias Zimmermann
BSI Business Systems Integration AG


«What is the difference
to the chess playing system
20 years ago?»
AlphaGo
Lee Sedol

AlphaGo Hardware
Powered by TPUs

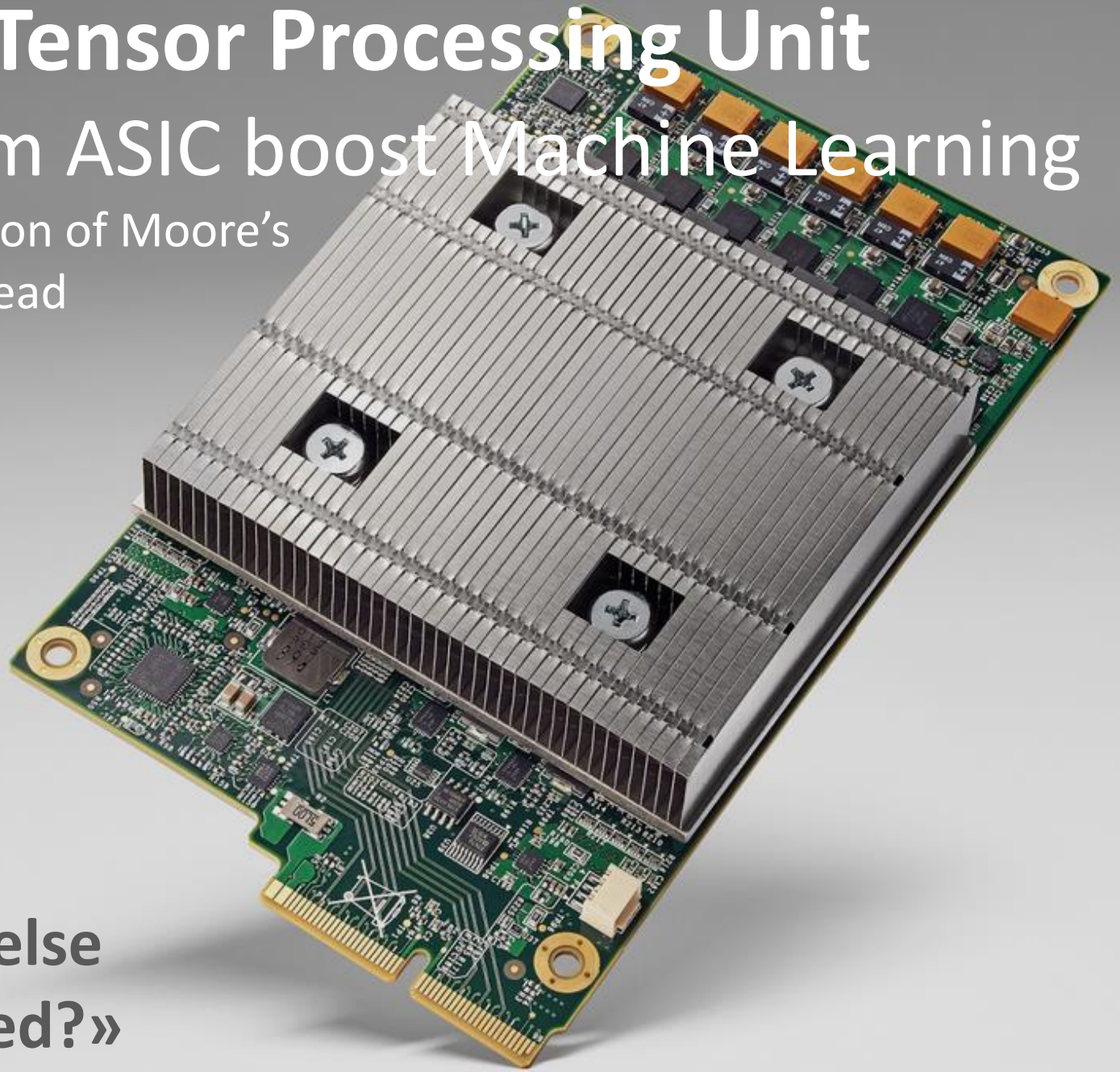
(GPUs are so yesterday ...)

TPU: Tensor Processing Unit

Custom ASIC boost Machine Learning

3 Generation of Moore’s

6 Years ahead



But …

«What else is needed?»

nature
International weekly journal of science

2014, DeepMind

NATURE | LETTER

日本語要約

Human-level control through deep reinforcement learning

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg & Demis Hassabis

Affiliations | Contributions | Corresponding authors

Nature **518**, 529–533 (26 February 2015) | doi:10.1038/nature14236

Received 10 July 2014 | Accepted 16 January 2015 | Published online 25 February 2015

Deep Reinforcement Learning

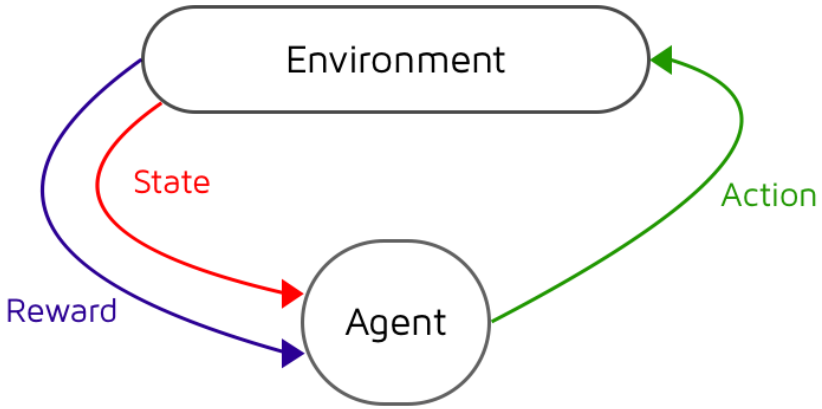
Markov Decision Process

➔ **Environment** (Atari Breakout)

➔ **Agent** performing **Actions** (Left, Right, Release Ball)

➔ **State** (Bricks, location / direction of ball, …)

➔ **Rewards** (A Brick is hit)



Deep Reinforcement Learning

Q-Learning (simplified)

➔ Markov Decision Process

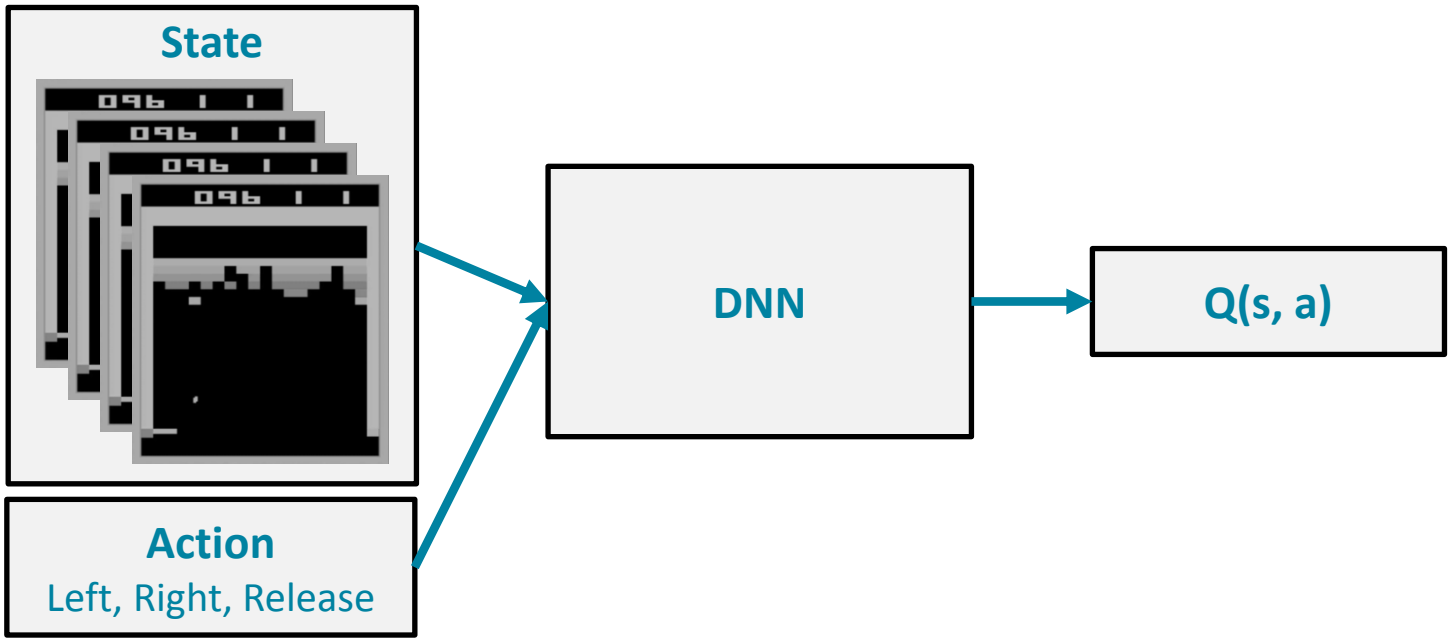
➔ **Q(s, a)** Highest sum of future **Rewards** for Action a

```
initialize Q randomly
assume initial state s_0
repeat
   execute a to maximize Q(s_i, a)
   observe r and new state s_i+1
   set Q = update(Q, r, s_i+1)
   set s_i = s_i+1
until terminated
```

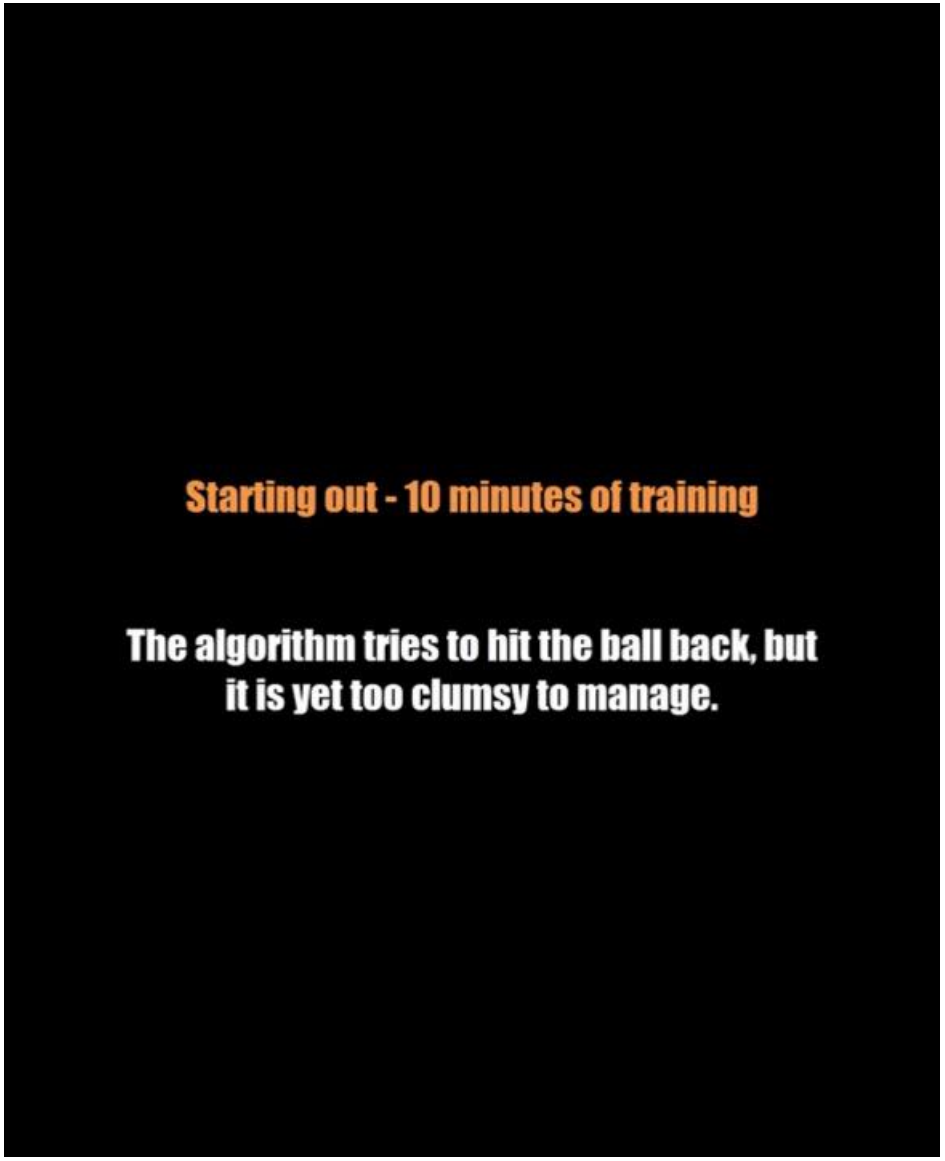
Deep Reinforcement Learning

Deep Q Learning (DQN)

- ➔ Q Learning
- ➔ Q(s, a) = **Deep Neural Network** (DNN)
- ➔ Retrain DNN regularly (using it's own experience)



DQN
Atari Breakout



Machine Learning Concepts

Data
Models
Training and Evaluation
ML Topics

Data

Models

Training and Evaluation

ML Topics

Getting the Data

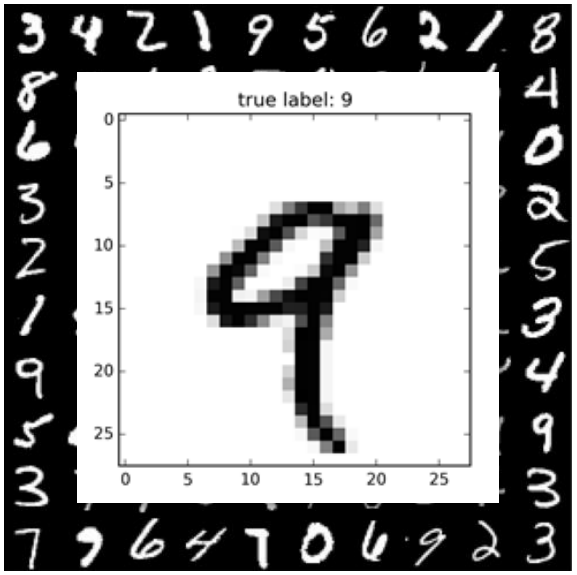
Challenges

- ➔ Getting the **RIGHT** data for the task
- ➔ And **LOTSs** of it
- ➔ However, there is never enough data …

Real World Lessons

- ➔ Crucial for successful ML projects
- ➔ Really not the exciting Part of ML
- ➔ Always takes much longer than planned

1245 **NEGATIVE** \
shallow , noisy and pretentious .
14575 **POSITIVE** \
one of the most splendid entertainments to emerge from the french film industry in years



86211,**B**,12.18,17.84,77.79, …
862261,**B**,9.787,19.94,62.11, …
862485,**B**,11.6,12.84,74.34, …
862548,**M**,14.42,19.77,94.48, …
862009,**B**,13.45,18.3,86.6, …

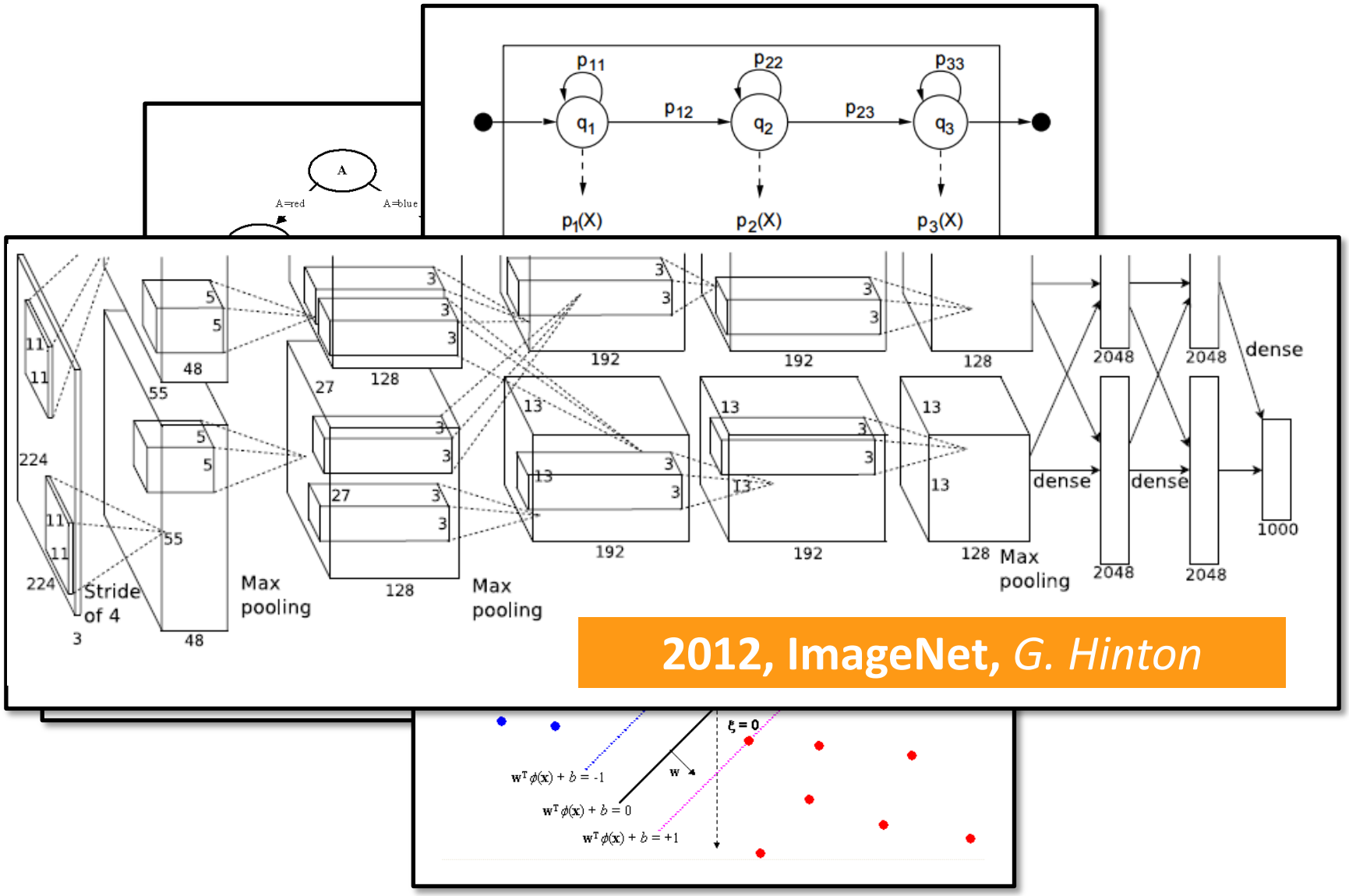
Rosemary, Rosmarinus officinalis

Data

Models

Training and Evaluation

ML Topics



2012, ImageNet, *G. Hinton*

Data
Models
Training and Evaluation
ML Topics

CS 229
Machine Learning
Handout #2: Course Schedule

Syllabus

- **Introduction** (1 class)
  Basic concepts.

- Supervised learning. (7 classes)
  Supervised learning setup. LMS.
  Logistic regression. Perceptron. Exponential family.
  Generative learning algorithms. Gaussian discriminant analysis.
  Support vector machines.
  Model selection and feature selection.
  Ensemble methods: Bagging, boosting.
  Evaluating and debugging learning algorithms.

- **Learning theory.** (3 classes)
  Bias/variance tradeoff. Union and Chernoff/Hoeffding bounds.
  VC dimension. Worst case (online) learning.
  Practical advice on how to use learning algorithms.

- Unsupervised learning. (5 classes)
  Clustering. K-means.
  EM. Mixture of Gaussians.
  Factor analysis.
  PCA (Principal components analysis).
  ICA (Independent components analysis).

- Reinforcement learning and control. (4 classes)
  MDPs. Bellman equations.
  Value iteration and policy iteration.
  Linear quadratic regulation (LQR). LQG.
  Q-learning. Value function approximation.
  Policy search. Reinforce. POMDPs.

Supervised Learning
- Learning from Examples
- Right Answers are known

Unsupervised Learning
- Discover Structure in Data
- Dimensionality Reduction

Reinforcement Learning
- Interaction with Dynamic Environment
- Reward Function

«Demo Time»

3 Demos

1. Outlier Detection (Brest Cancer Data)
2. Sentiment Analysis (Movie Reviews)
3. Pattern Recognition (MNIST Database)

Outlier Detection
Diagnostic Breast Cancer Data

WDBC Data Set

«12.86,18,83.19,506.3,0.09934, … »

➔ Malignant or Benign?

86211,**B**,12.18,17.84,77.79, …
862261,**B**,9.787,19.94,62.11, …
862485,**B**,11.6,12.84,74.34, …
862548,**M**,14.42,19.77,94.48, …
862009,**B**,13.45,18.3,86.6, …

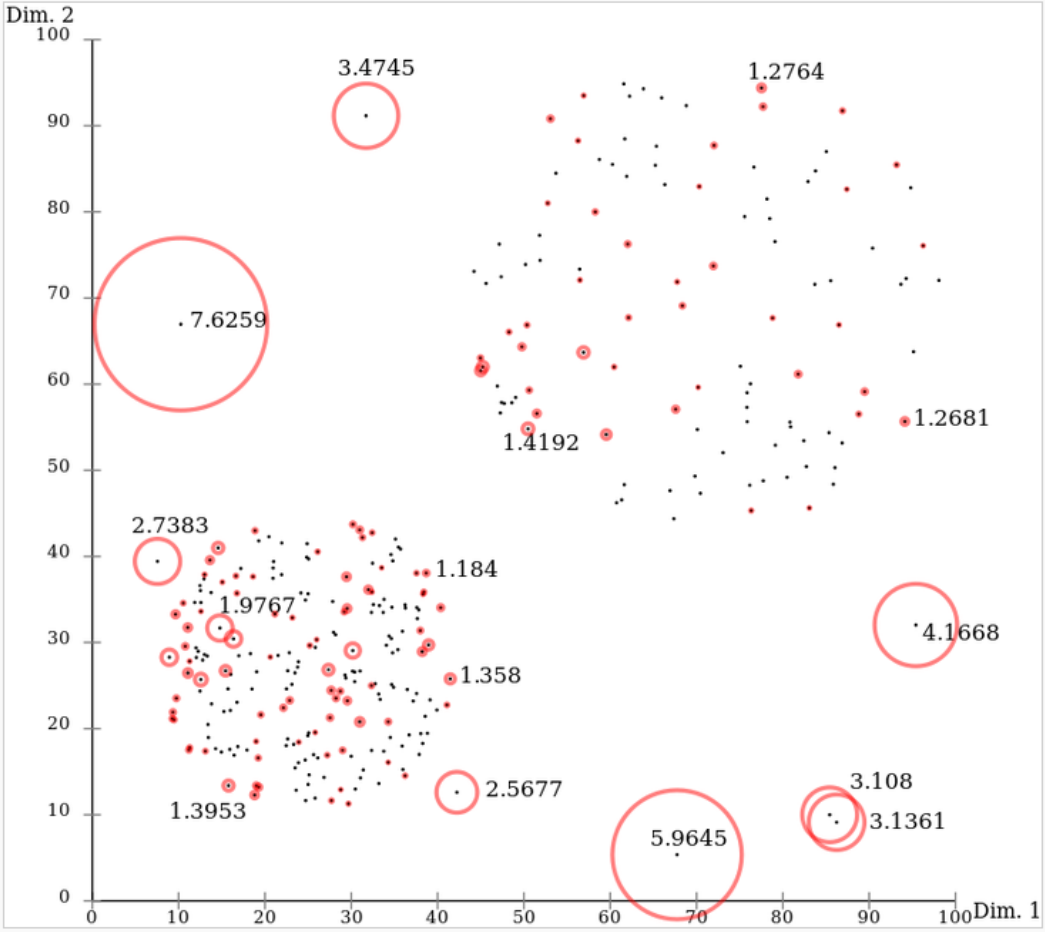
Model/Algorithm

➔ LOF (Local Outlier Factor)

Direct Java Implementation

➔ Java

Local Outlier Factor



LOF scores as visualized by ELKI. While the upper right cluster has a comparable density to the outliers close to the bottom left cluster, they are detected correctly.

Natural Language Processing
Sentiment Analysis

Movie Review Data

«please, someone, stop eric schaeffer before he makes another film»

➔ Positive or negative?

1245 **NEGATIVE** \
shallow , noisy and pretentious .
14575 **POSITIVE** \
one of the most splendid entertainments to emerge from the french film industry in years

Model

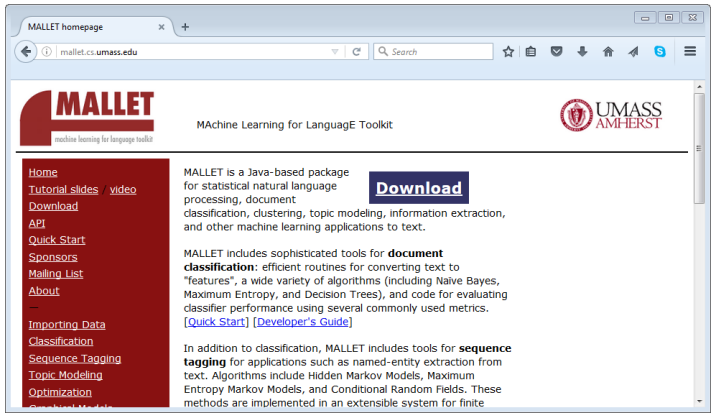
➔ Naive Bayes

Mallet Toolkit

➔ Statistical NLP

➔ Open Source (CPL)

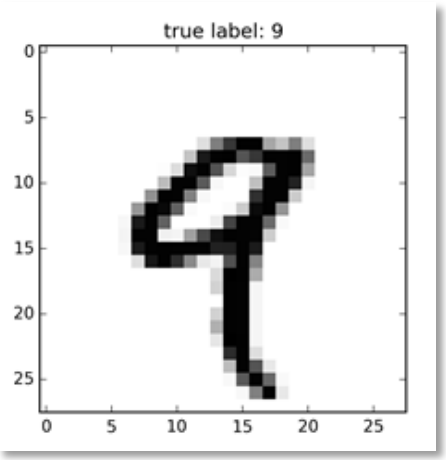
➔ Java



Pattern Recognition
Handwritten Digits

Handwritten Digits

- MNIST database
- Which digit is this?

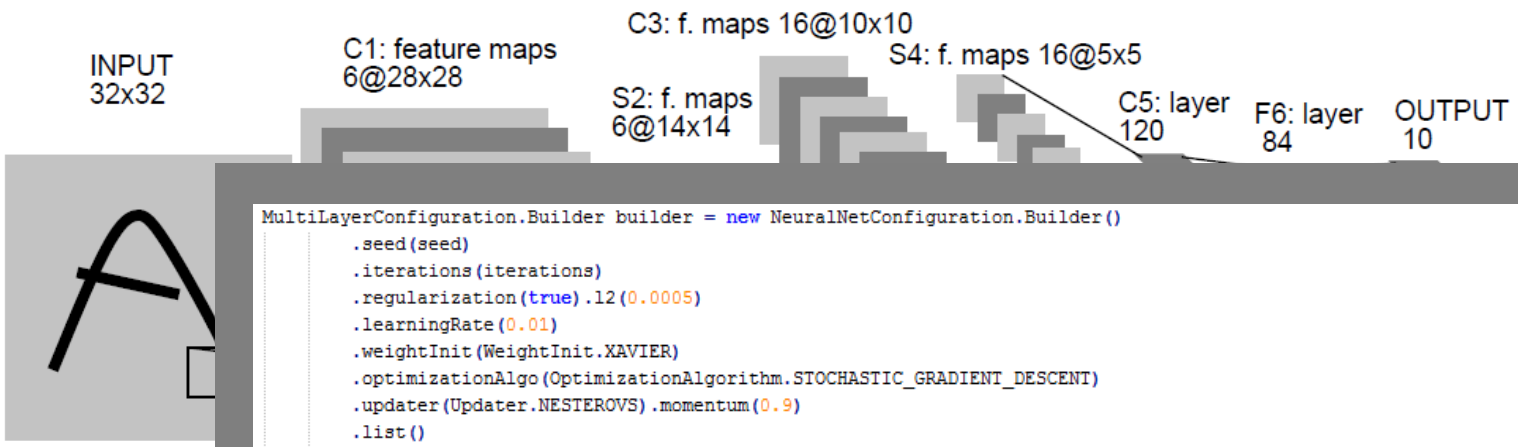


Model

- Deep Neural Network (LeNet-5)

Deeplearning4j

- Deep Learning Library
- Open Source (Apache)
- Java




```
MultiLayerConfiguration.Builder builder = new NeuralNetConfiguration.Builder()
        .seed(seed)
        .iterations(iterations)
        .regularization(true).l2(0.0005)
        .learningRate(0.01)
        .weightInit(WeightInit.XAVIER)
        .optimizationAlgo(OptimizationAlgorithm.STOCHASTIC_GRADIENT_DESCENT)
        .updater(Updater.NESTEROVS).momentum(0.9)
        .list()
        .layer(0, new ConvolutionLayer.Builder(5, 5)
                //nIn and nOut specify depth. nIn here is the nChannels and nOut is the number of filters to be applied
                .nIn(nChannels)
                .stride(1, 1)
                .nOut(20)
                .activation("identity")
                .build())
        .layer(1, new SubsamplingLayer.Builder(SubsamplingLayer.PoolingType.MAX)
                .kernelSize(2,2)
                .stride(2,2)
                .build())
        .layer(2, new ConvolutionLayer.Builder(5, 5)
                //Note that nIn need not be specified in later layers
                .stride(1, 1)
                .nOut(50)
                .activation("identity")
                .build())
        .layer(3, new SubsamplingLayer.Builder(SubsamplingLayer.PoolingType.MAX)
                .kernelSize(2,2)
                .stride(2,2)
                .build())
        .layer(4, new DenseLayer.Builder().activation("relu")
                .nOut(500).build())
        .layer(5, new OutputLayer.Builder(LossFunctions.LossFunction.NEGATIVELOGLIKELIHOOD)
                .nOut(outputNum)
                .activation("softmax")
                .build())
        .setInputType(InputType.convolutionalFlat(28,28,1)) //See note below
        .backprop(true)
        .pretrain(false);
```

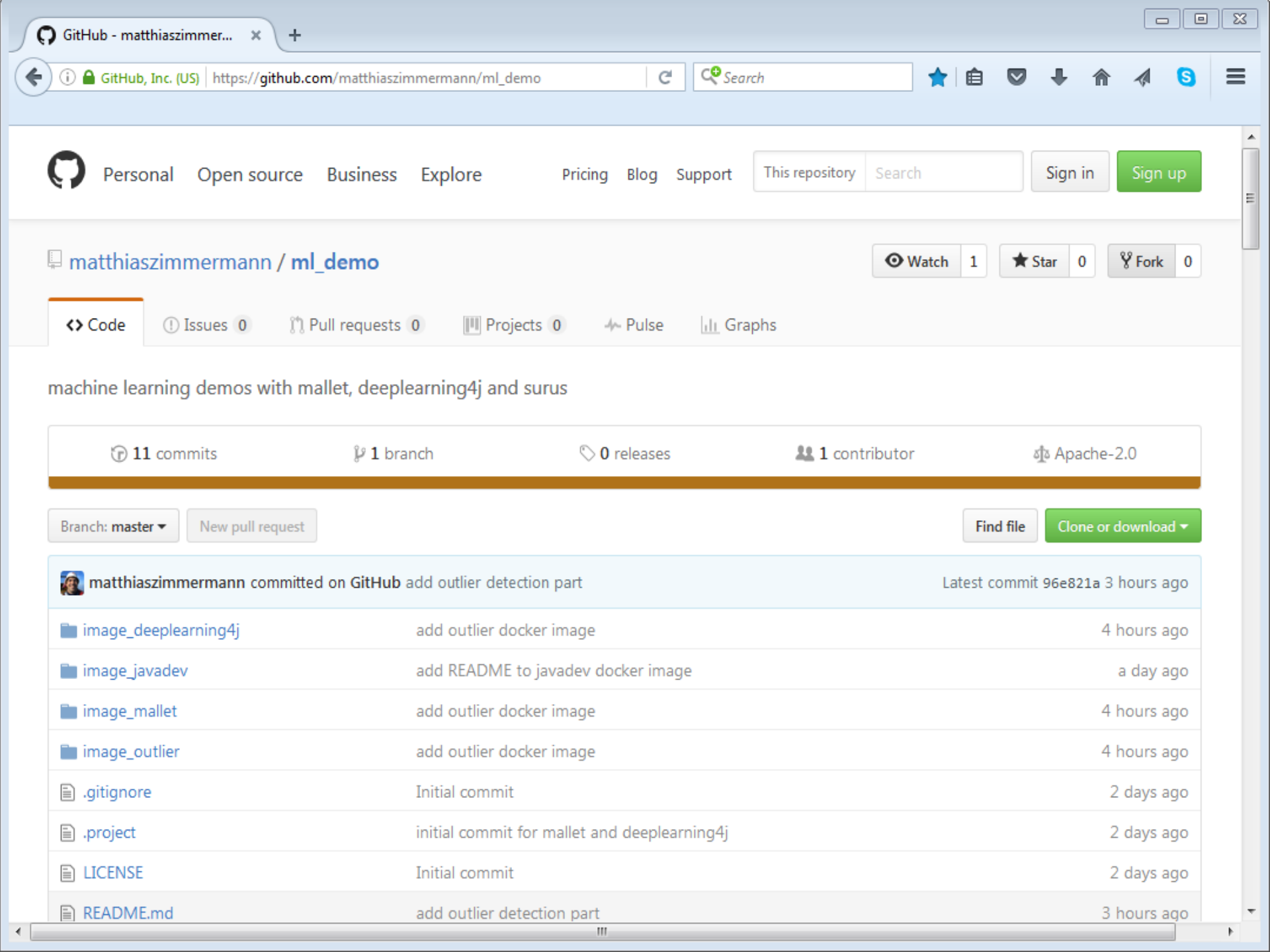
n connections

Fig. 2. Architec

e. a set of units whose weigh

1998,

LeCun


GitHub - matthiaszimmer...
GitHub, Inc. (US) https://github.com/matthiaszimmermann/ml_demo
Search
Personal Open source Business Explore Pricing Blog Support
This repository Search
Sign in Sign up
matthiaszimmermann / ml_demo
Watch 1 Star 0 Fork 0
Code Issues 0 Pull requests 0 Projects 0 Pulse Graphs
machine learning demos with mallet, deeplearning4j and surus
11 commits 1 branch 0 releases 1 contributor Apache-2.0
Branch: master New pull request
Find file Clone or download
matthiaszimmermann committed on GitHub add outlier detection part
Latest commit 96e821a 3 hours ago
image_deeplearning4j add outlier docker image 4 hours ago
image_javadev add README to javadev docker image a day ago
image_mallet add outlier docker image 4 hours ago
image_outlier add outlier docker image 4 hours ago
.gitignore Initial commit 2 days ago
.project initial commit for mallet and deeplearning4j 2 days ago
LICENSE Initial commit 2 days ago
README.md add outlier detection part 3 hours ago

Recent Advances

Recent Advances

1. DeepFace (2014, FaceBook, 2014)
2. Automatic Image Captioning (2014, Stanford, …)
3. TensorFlow Library (2015, Google)
4. DNC (2016, Google DeepMind)

DeepFace: Closing the Gap to Human-Level Performance in Face Verification

Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, Lior Wolf

Conference on Computer Vision and Pattern Recognition (CVPR) · June 24, 2014

Applied Machine Learning, Machine Learning, Facebook AI Research (FAIR), Artificial Intelligence

2014, FaceBook

Abstract

In modern face recognition, the conventional pipeline consists of four stages: detect => align => represent => classify. We revisit both the alignment step and the representation step by employing explicit 3D face modeling in order to apply a piecewise affine transformation, and derive a face representation from a nine-layer deep neural network. This deep network involves more than 120 million parameters using several locally connected layers without weight sharing, rather than the standard convolutional layers. Thus we trained it on the largest facial dataset to-date, an identity labeled dataset of four million facial images belonging to more than 4,000 identities.

The learned representations coupling the accurate model-based alignment with the large facial database generalize remarkably well to faces in unconstrained environments, even with a simple classifier. Our method reaches an accuracy of 97.35% on the Labeled Faces in the Wild (LFW) dataset, reducing the error of the current state of the art by more than 27%, closely approaching human-level performance.

Download Paper

Related Blog Posts

ECCV Brings Together the Brightest Minds in Computer Vision
by Kelly Berschauer
Oct 8, 2016

Facebook Hosts Distinguished Faculty
by Kelly Berschauer
Sep 29, 2016

Facebook to Accelerate Global AI Research with New GPU Program Recipients
by Ari Entin
Aug 29, 2016

Learning to Segment
by Piotr Dollar
Aug 25, 2016

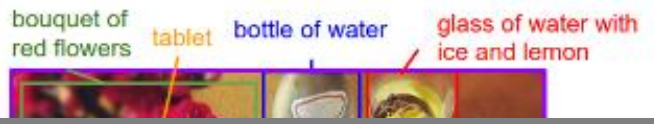
fastText

2014, Stanford

Deep Visual-Semantic Alignments for Generating Image Descriptions

Andrej Karpathy Li Fei-Fei
Department of Computer Science, Stanford University
{karpathy,feifeili}@cs.stanford.edu

Abstract



construction worker in orange safety vest is working on road.

two young girls are playing with lego toy.

boy is doing backflip on wakeboard.

TensorFlow ™

GET STARTED TUTORIALS HOW TO MOBILE API RESOURCES ABOUT

Fork me on GitHub

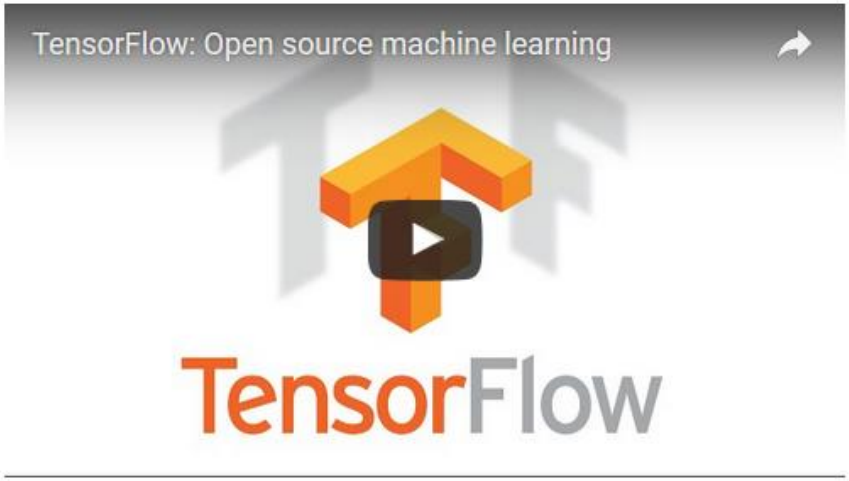
TensorFlow is an Open Source Software Library for Machine Intelligence

GET STARTED

2015, Google

About TensorFlow

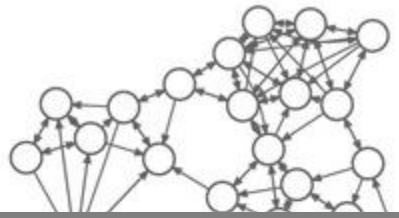
TensorFlow™ is an open source software library for numerical computation using data flow graphs. Nodes in the graph represent mathematical operations, while the graph edges represent the multidimensional data arrays (tensors) communicated between them. The flexible architecture allows you to deploy computation to one or more CPUs or GPUs in a desktop, server, or mobile device with a single API. TensorFlow was originally developed by researchers and engineers working on the Google Brain Team within Google's Machine Intelligence research organization for the purposes of conducting machine learning and deep neural networks research, but the system is general enough to be



Diff

Mind

Random Training Graph



London Underground

We aim to combine the advantages of neural and computational processing by providing a neural network with read–write access to external memory.

The whole system is differentiable, and can therefore be trained end-to-end with gradient descent, allowing the network to learn how to operate and organize the memory in a goal-directed manner.

(OxfordCircus, PiccadillyCircus, Bakerloo)
(OxfordCircus, NottingHillGate, Central)
(OxfordCircus, Euston, Victoria)

- 84 edges in total

...
(Westminster, GreenPark, Jubilee)
(GreenPark, BondSt, Jubilee)

(Holborn, LeicesterSq, Piccadilly)
(LeicesterSq, PiccadillyCircus, Piccadilly)

In a *k*
calle
ansu *s,*
fami *ve a*
block puzz

Plato liker , would
remain fix minds
can be shaped and reshaped by experience. But the wax of our memories does not just

Food for Thought

Artificial Intelligence and Robotics + Add to myFT

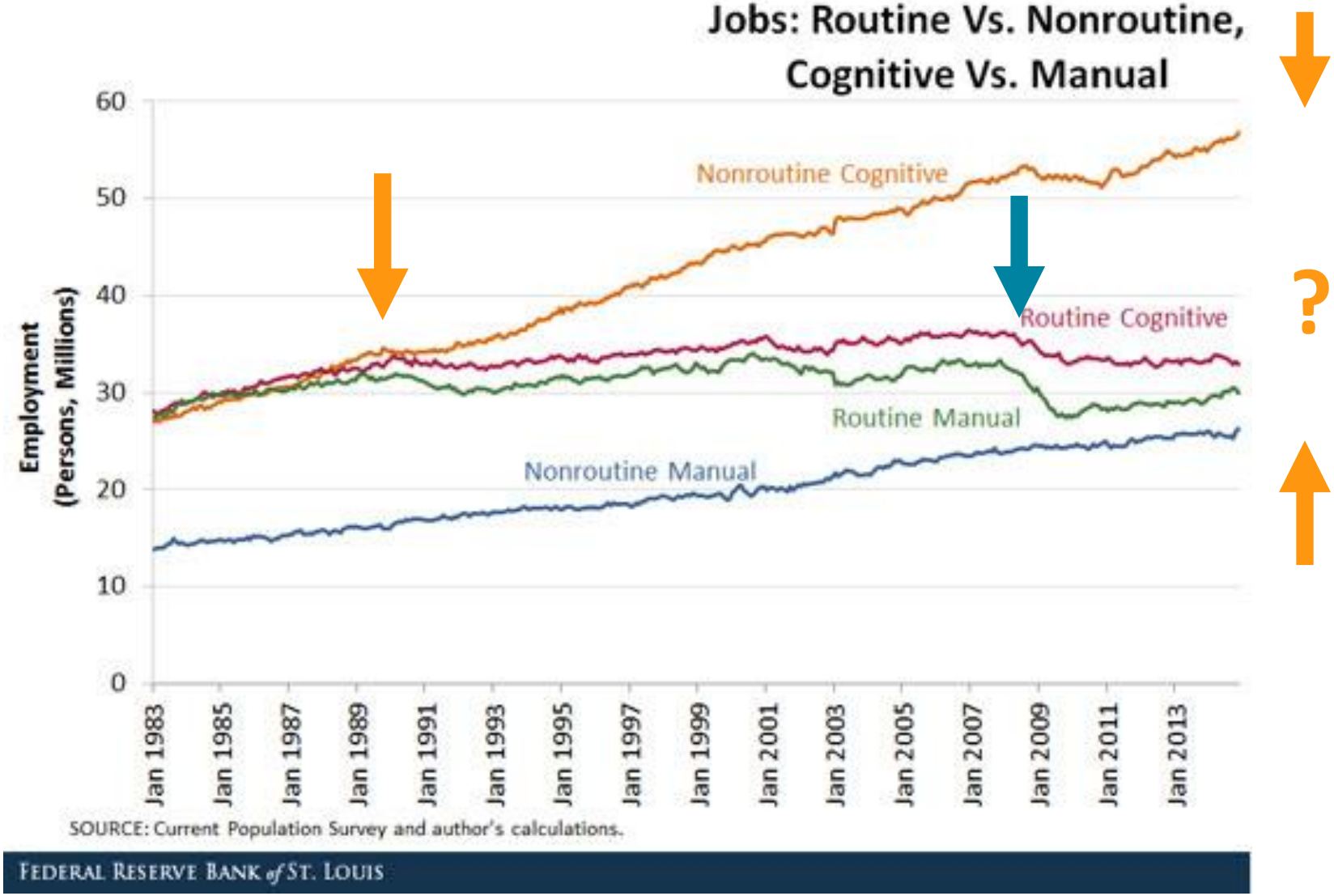
AI and robots threaten to unleash mass unemployment, scientists warn

Intelligent machines will soon replace human workers in all sectors of economy

Read latest:

Market grows for 'regtech', or AI for regulation

OCTOBER 14, 2016


Jobs: Routine Vs. Nonroutine, Cognitive Vs. Manual
Employment (Persons, Millions)
60
50
40
30
20
10
0
Nonroutine Cognitive
Routine Cognitive
Routine Manual
Nonroutine Manual
?
Jan 1983
Jan 1985
Jan 1987
Jan 1989
Jan 1991
Jan 1993
Jan 1995
Jan 1997
Jan 1999
Jan 2001
Jan 2003
Jan 2005
Jan 2007
Jan 2009
Jan 2011
Jan 2013
SOURCE: Current Population Survey and author's calculations.
FEDERAL RESERVE BANK of ST. LOUIS

The positive Side

Statement by Lee Sedol

Lee replied that playing against the machine had rekindled his passion for Go. As with Fan Hui, AlphaGo had opened his eyes to a new side of the game. “I have improved already,” Lee said. “It has given me new ideas.” He has not lost a match since.

Take Home Messages

What have we seen?

- ➔ Deep Learning was a break-through for ML
- ➔ ML now reaches human level performance on many tasks

Future

- ➔ Some jobs will get obsolete
- ➔ New types of jobs will most likely emerge
- ➔ Transition will be challenging

Get your Hands dirty ☺

- ➔ Tools simple enough to use outside Labs
- ➔ Examples on Github to help Newbies

Thanks

@ZimMatthias

Evaluate the Sessions

Sign in and vote at **eclipsecon.org**

- 1 0 + 1