

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

Deep learning: algorithms and deployment

iSense Big Data Summit

Rob Romijnders

robromijnders.github.io

RomijndersRob@gmail.com

October 24, 2017

Overview

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

1 Why

2 How

3 What

4 Deploy

5 Wrap up

6 Questions

Object detection

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

Faster R-CNN

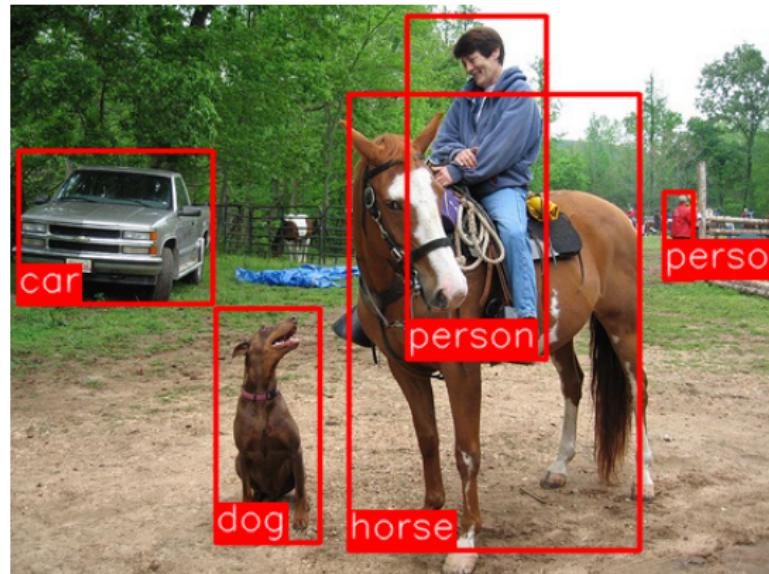


Figure: faster rcnn (image from Github *mitmul*)

Apple siri

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions



Siri



Figure: Photo: cultofmac.com

GMail reply

Why

Deploy

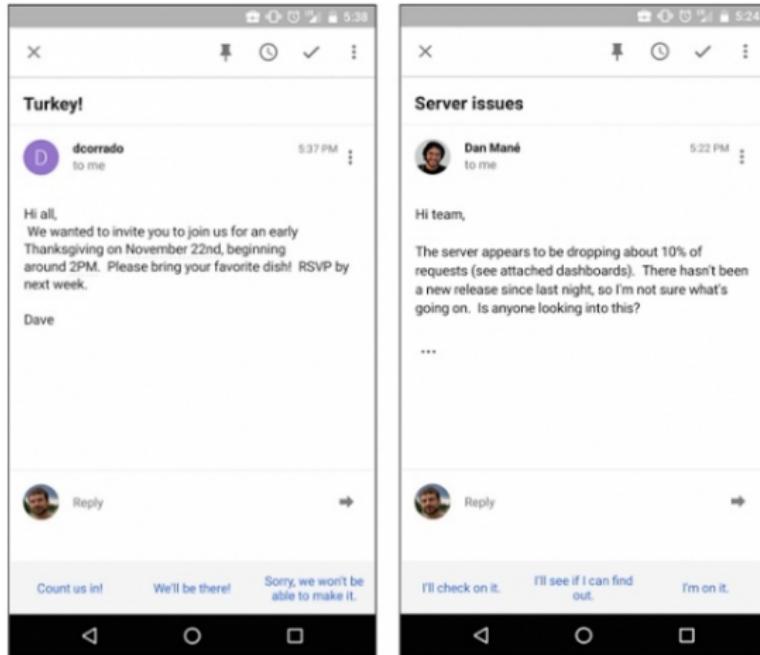


Figure: Photo: Greg Corrado, Google Research Blog

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions



Google Research Blog

The latest news from Research at Google

A Neural Network for Machine Translation, at Production Scale

Tuesday, September 27, 2016

Figure: credits: Google research blog

Segmentation

October 24,
2017

Rob
Romijnders

Why
How
What
Deploy
Wrap up
Questions



object classes	building	grass	tree	cow	sheep	sky	airplane	water	face	car
bicycle	flower	sign	bird	book	chair	road	cat	dog	body	boat

Figure: Semantic Segmentation with CNN (image from jamie.shotton.org) or see [youtube.com/watch?v=Nok6XludcQ](https://www.youtube.com/watch?v=Nok6XludcQ)

Pose estimation

Why

Deploy

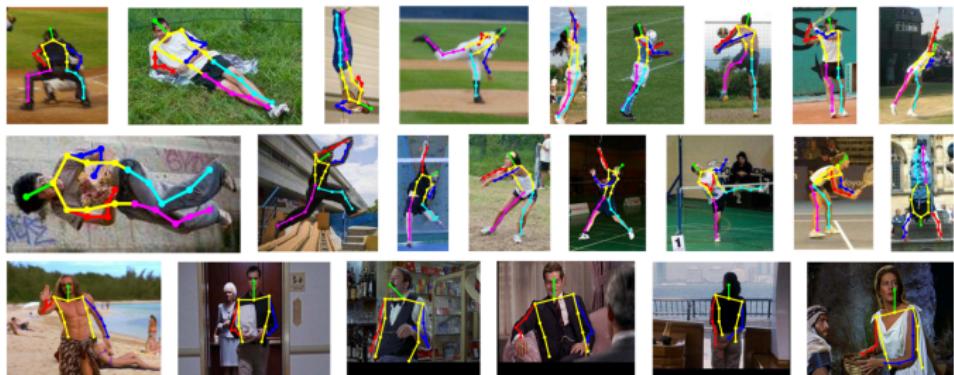


Figure: Source

October 24,
2017

Rob
Romijnders

Why
How
What
Deploy
Wrap up
Questions



Figure: Credits:
popularmechanics.com/technology/a19863/googles-alphago-ai-wins-second-game-go/

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

Generative modelling

Speech generation

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions



Figure: Source: deepmind.com/blog/wavenet-launches-google-assistant/

Generation

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

THE MULTIVERSE —

Movie written by algorithm turns out to be hilarious and intense

For *Sunspring*'s exclusive debut on Ars, we talked to the filmmakers about collaborating with an AI.

ANNALEE NEWITZ - 6/9/2016, 12:30 PM



Sunspring, a short science fiction movie written entirely by AI, debuts exclusively on Ars today.

Figure: <http://arstechnica.com/the-multiverse/2016/06/an-ai-wrote-this-movie-and-its-strangely-moving/>

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

Why not

Which data not

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

YES: media type data

- 1 Text, language, speech
- 2 Images, video, maps
- 3 Time-series, stocks, valuta
- 4 Games, robots

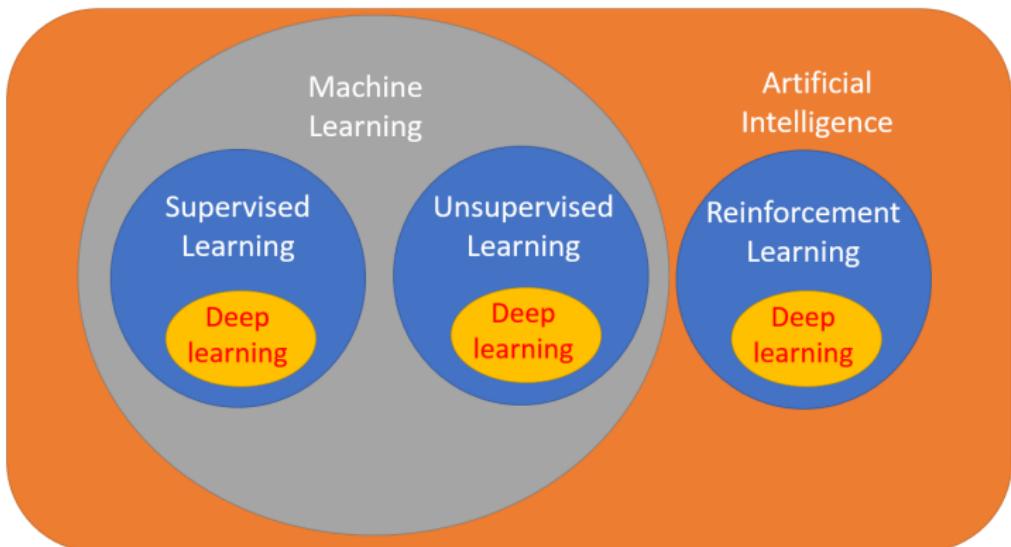
NO: categorical data

- 1 Properties
- 2 Features
- 3 Categories

Overview of the field

Why

How



October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

Neural Networks

Neural nets

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

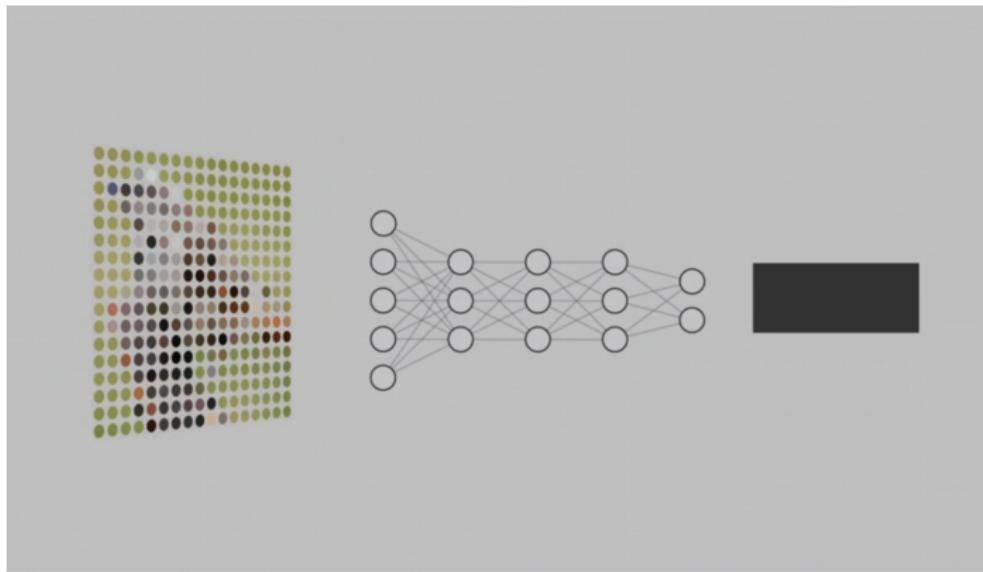


Figure: Neural network explained (credits: Blaise Aguera y Arcas)

Basic equation

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

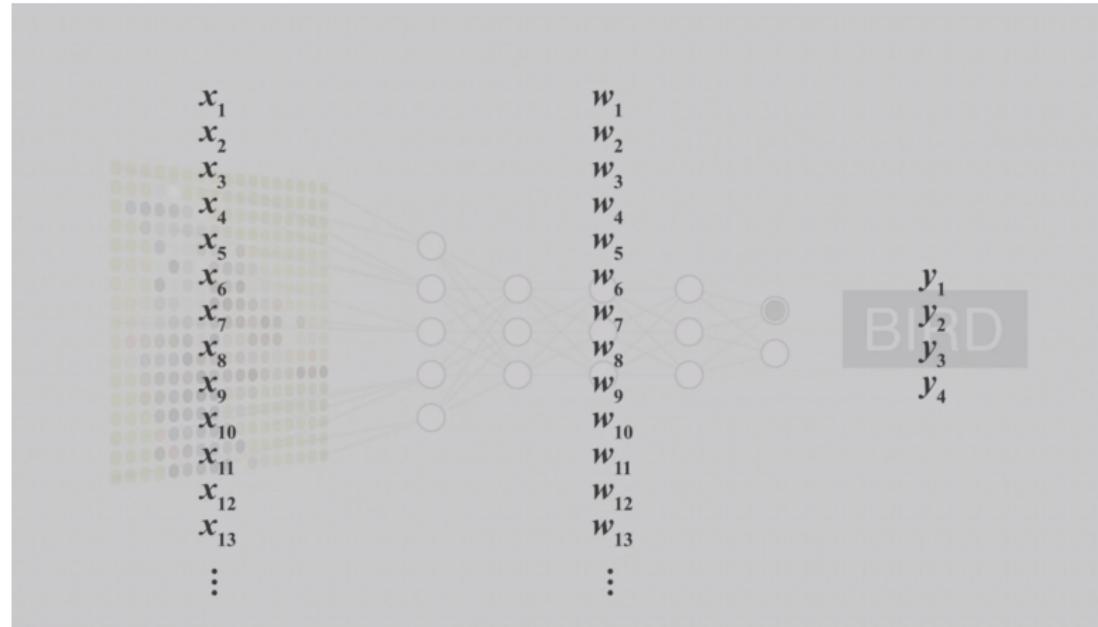


Figure: Neural net in algebraic form (credits: Blaise Aguera y Arcas)

Template equation

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions



Figure: Template equation neural net (credits: Blaise Aguera y Arcas)

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

How to use

How to use them

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

$$W \cdot X = y$$

$$2 \cdot 3 = y$$

Forward inference

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

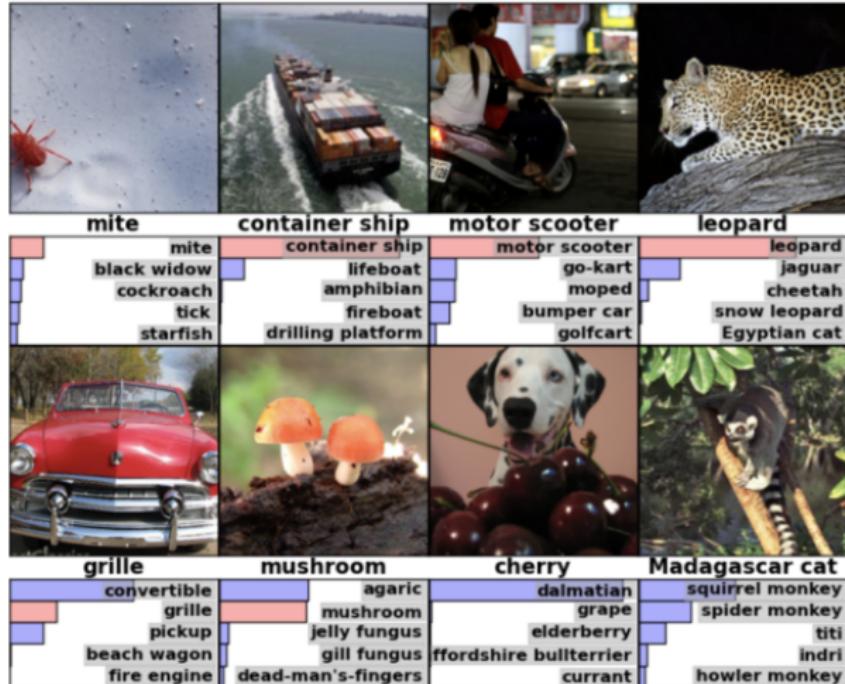


Figure: Forward inference CNN (Krizhevsky et al. 2012)

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

How to train

How to train them

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

$$(x, y) = (3, 6)$$

$$w \cdot x = y$$

$$w \cdot 3 = 6$$

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

$$w \cdot x \approx y$$

$$y - w \cdot x \approx 0$$

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

$$\text{Loss} = (y - w \cdot x)^2 \implies 0$$

Gradient Descent

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

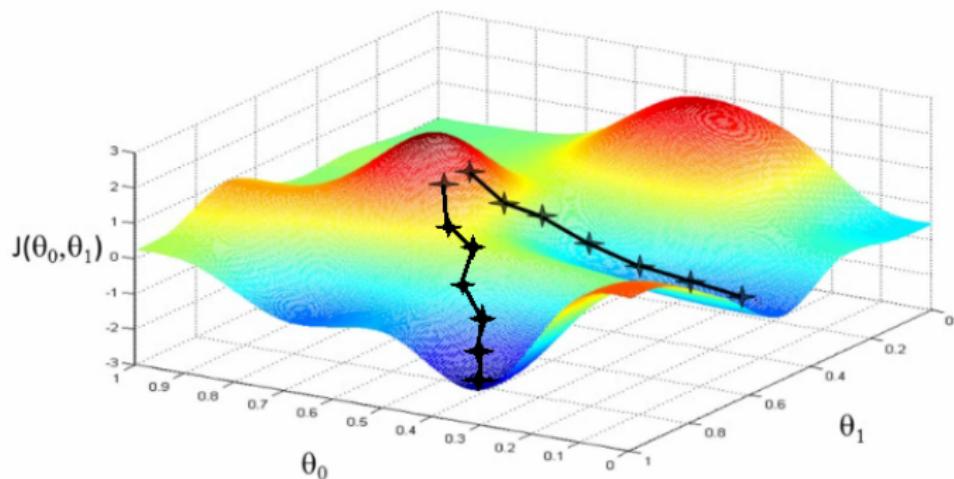


Figure: Credits: ML blog Vasilis Vryniotis

Loss function

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

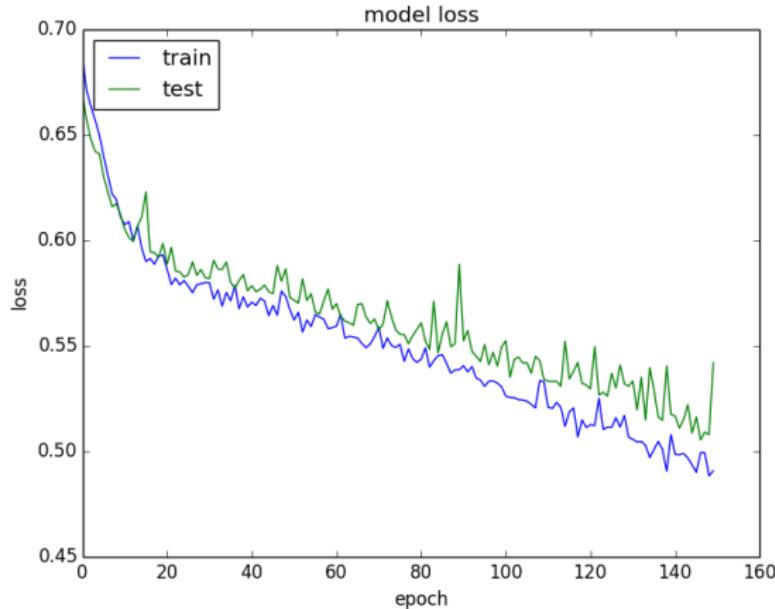


Figure: (<http://machinelearningmastery.com/>)

Training a neural network

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

- 1 Define a neural network
- 2 Define the loss function
- 3 Do backpropagation

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

1. Define a neural network

One hidden layer

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions

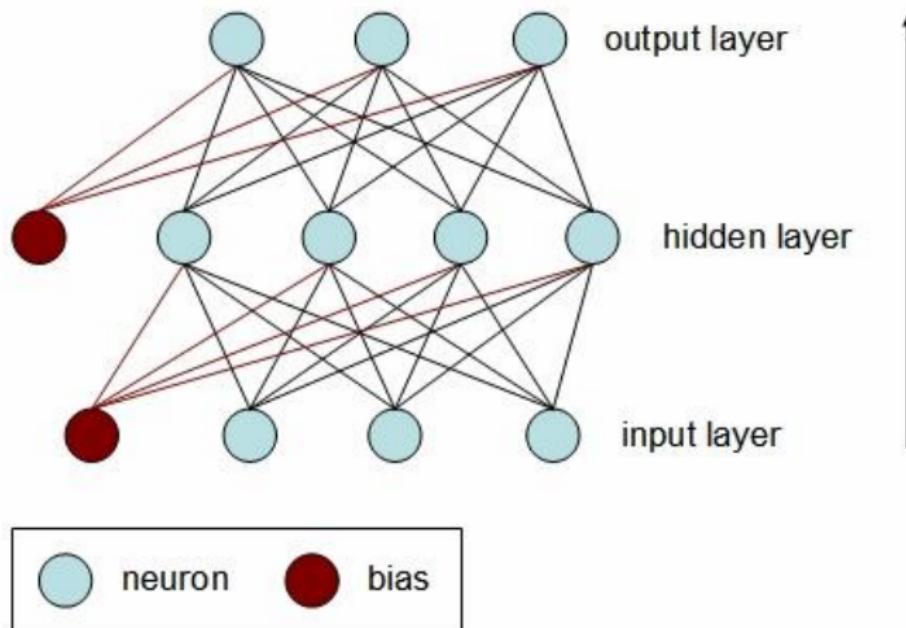


Figure: Neural network with one hidden layer

Activation function

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions

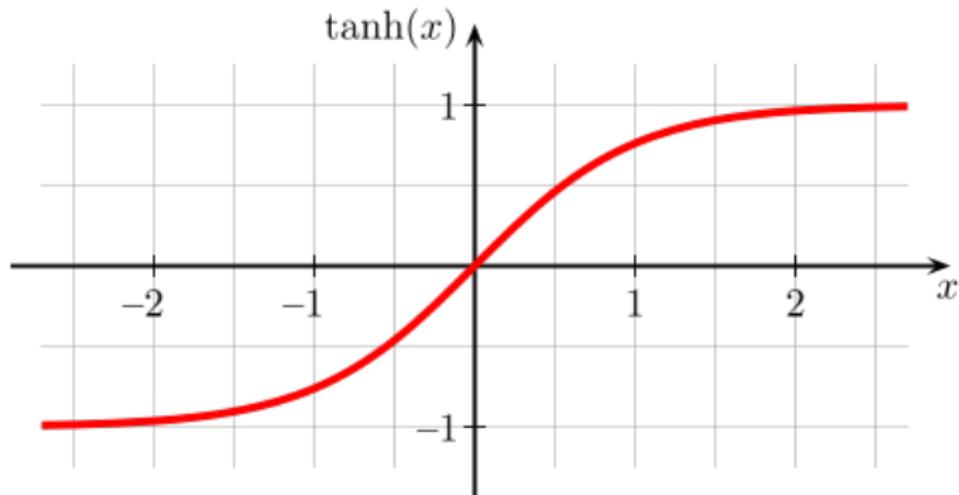


Figure: Hyperbolic tangent as activation function

Python code

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions

```
#FF
def step(x):
    h = np.tanh(np.dot(x, W1) + b1)
    y = np.dot(h, W2) + b2
    return y

y = step(x)
```

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison

What
Deploy

Wrap up

Questions

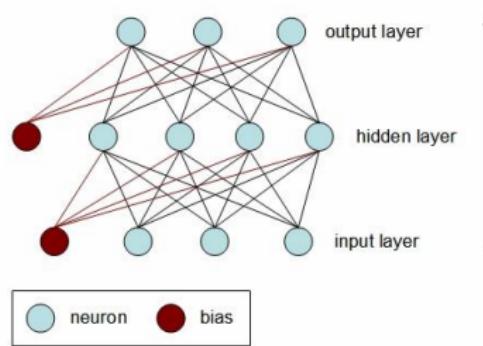


Figure: Neural network with one hidden layer

$W1.\text{shape} \Rightarrow (3, 4)$
 $b1.\text{shape} \Rightarrow (4,)$
 $W2.\text{shape} \Rightarrow (4, 3)$
 $b2.\text{shape} \Rightarrow (3,)$

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

2. Define the loss function

Regression

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

$$\text{Loss} = (y_{correct} - \hat{y})^2$$

Classification

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

$$Loss = - \sum_i y_{correct} \log(\hat{y})$$

Softmax

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions

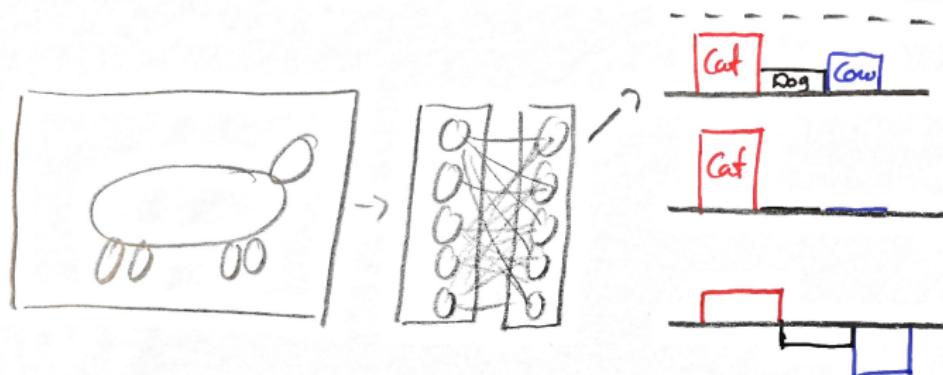


Figure: Gradient from softmax

Equations

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

1 Softmax

$$\text{softmax}(a)_i = \frac{e^{a_i}}{\sum_j e^{a_j}}$$

2 Cross entropy

$$\text{loss} = \text{cross}(p, q) = - \sum_j p_j \log(q_j) = -\log(q_{y_{\text{correct}}})$$

3 Gradient

$$\frac{\delta \text{Loss}}{\delta a_i} = q_i - p_i$$

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

3. Do stochastic gradient descent

Gradient Descent

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions

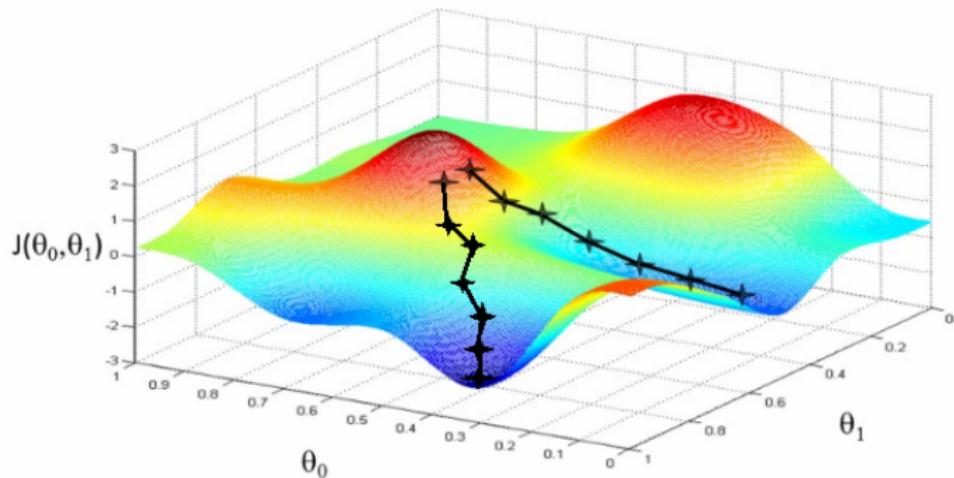


Figure: Credits: ML blog Vasilis Vryniotis

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions

Minimize loss

$$w = w - \eta \frac{\delta Loss}{\delta w}$$

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

$$\frac{\delta Loss}{\delta w} = \frac{\delta Loss}{\delta a} \frac{\delta a}{\delta h_l} \frac{\delta h_l}{\delta h_{l-1}} \frac{\delta h_{l-1}}{\delta h_{l-2}} \dots \frac{\delta h_1}{\delta h_0} \frac{\delta h_0}{\delta w}$$

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

Stochastic gradient descent

Use small batches of size e.g. 32

Two problems

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

1 Too many weights

$$3 \times 1024 \times 100 + 100 \times 3 = 307500$$

2 No domain knowledge

Two architectures

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

1 Convolution

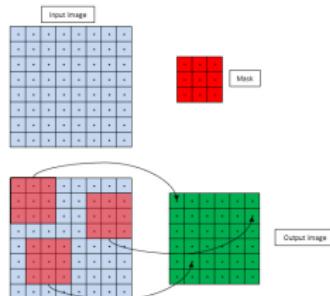


Figure: developer.amd.com

2 Recursion

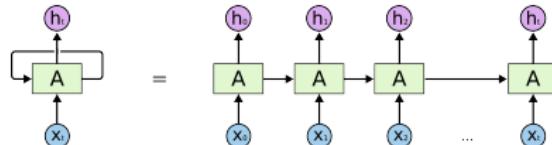


Figure: credits: colah.github.io

Detect patterns

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions

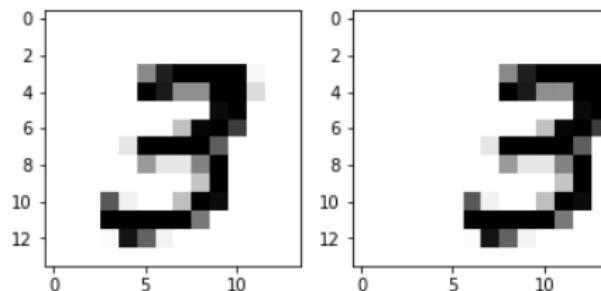


Figure: Translation of image does not change class

Convolution

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison

What
Deploy
Wrap up
Questions

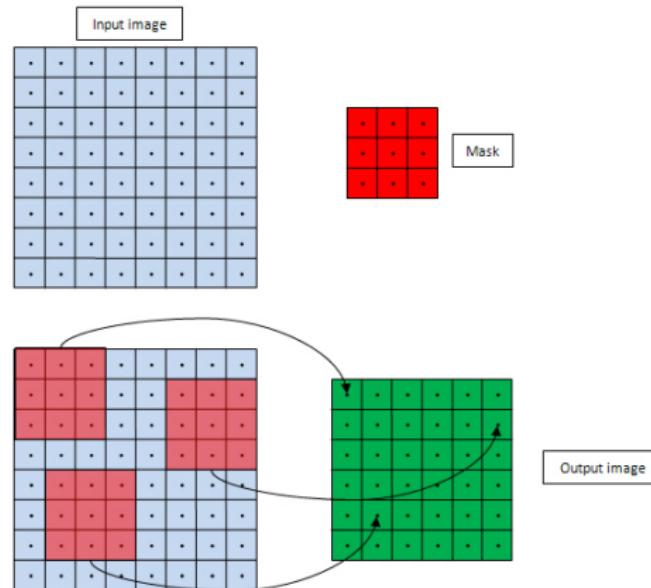


Figure: developer.amd.com

Python code

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions

#CNN

```
def step(x):
    h = np.tanh(convolve(x, W1) + b1)
    y = np.dot(h.flatten(), W2) + b2
    return y

y = step(x)
```

CNN

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

ILSVRC top-5 error on ImageNet

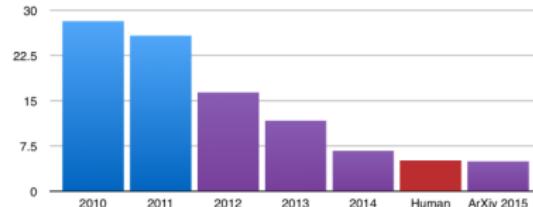


Figure: ImageNet results (credits jackkelly.github.io/)

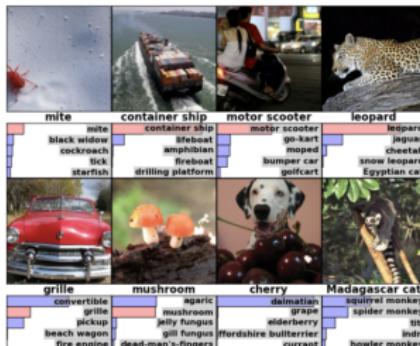


Figure: Forward inference CNN (Krizhevsky et al. 2012)

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions

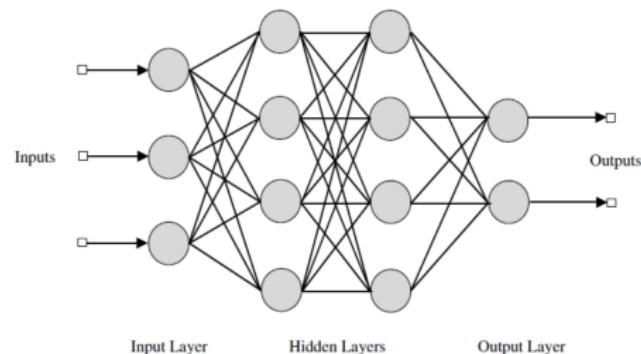


Figure: Feedforward neural network

$$3 \times 1024 \times 100 + 100 \times 3 = 307500$$

$$3 \cdot 5 \cdot 5 \cdot 100 + 100 \times 3 = 7800$$

Recursion

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

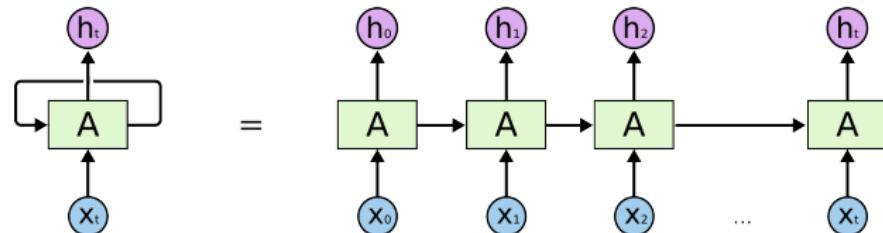


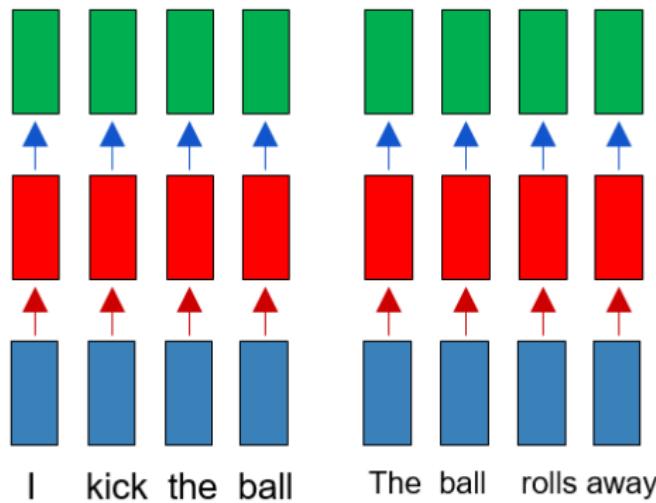
Figure: credits: colah.github.io

Word tagging

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions



Connect over time

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions

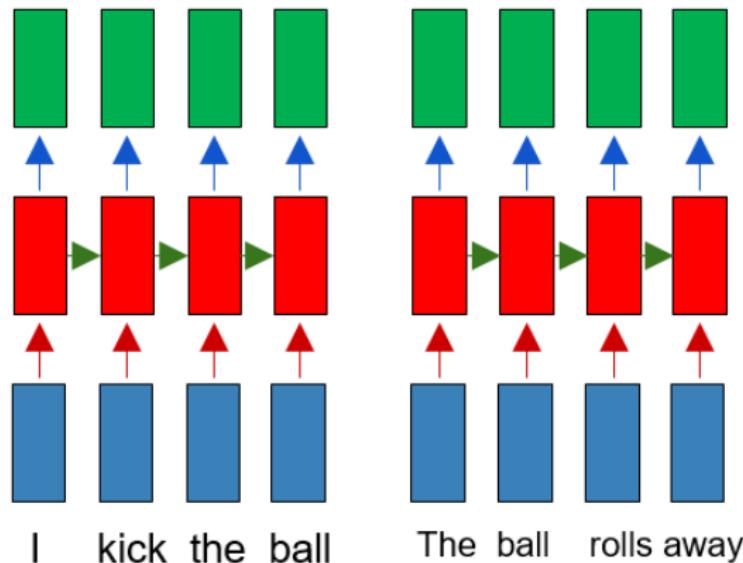


Figure: [RNN in 112 lines](#)

Python code

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions

```
#FF
def step(x):
    h = np.tanh(np.dot(x, W1) + b1)
    y = np.dot(h, W2) + b2
    return y

y = step(x)
```

Python code

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

```
# RNN
rnn = RNN()

class RNN:
    def step(self, x):
        self.h = np.tanh(np.dot(self.h, W1)
                        + np.dot(x, W2))
        y = np.dot(self.h, W3)
        return y

y = rnn.step(x)
```

Python code

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy

Wrap up

Questions

```
#FF
def step(x):
    h = np.tanh(np.dot(W1, x) + b1)
    y = np.dot(W2, h) + b2
    return y

#RNN
class RNN:
    def step(self, x):
        self.h = np.tanh(np.dot(W1, self.h)
                         + np.dot(W2, x))
        y = np.dot(W3, self.h)
    return y
```

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

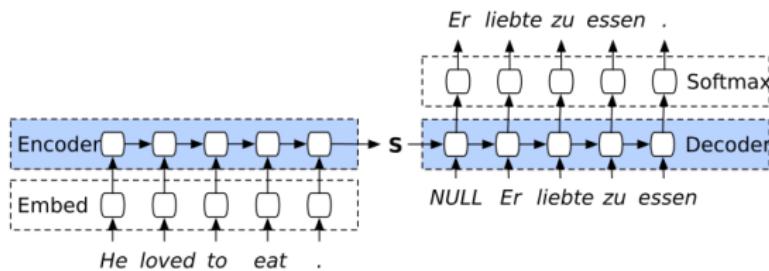


Figure: RNN for Machine Translation (credits: smerity.com)

Long Short-term memory

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c)$$

$$h_t = o_t \tanh(c_t)$$

LSTM

October 24,
2017

Rob
Romijnders

Why
How
FFNN
CNN
RNN
Comparison
What
Deploy
Wrap up
Questions

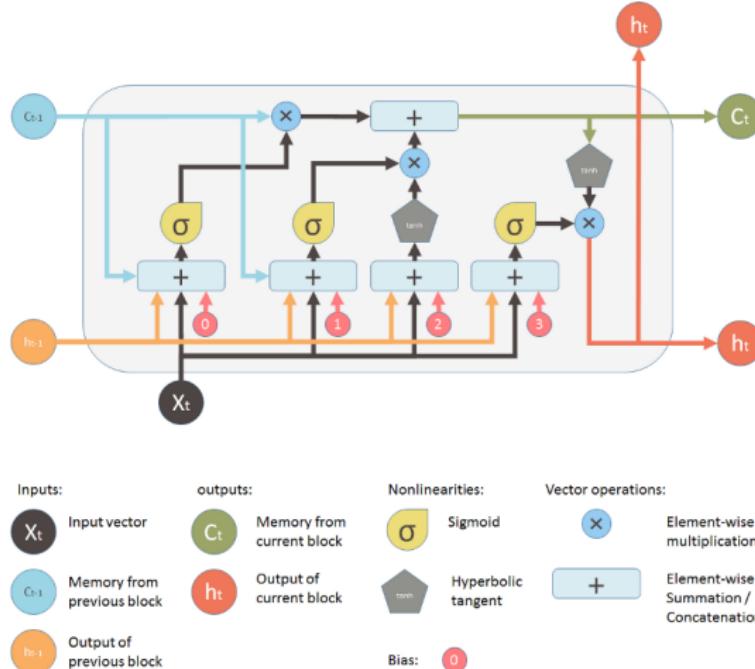


Figure: Diagram depicting LSTM block

Author: Shi Yan. Source: medium.com/@shiyanshiyan/

Span

October 24,
2017

Rob
Romijnders

Why

How

FFNN

CNN

RNN

Comparison

What

Deploy

Wrap up

Questions

1 CNN

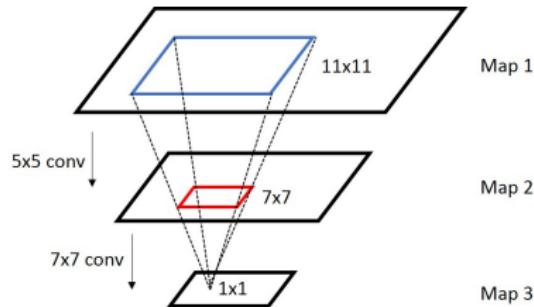


Figure: credits: cvmarcher.com/

2 RNN

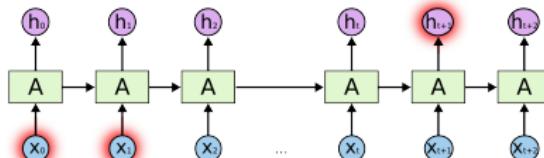


Figure: credits colah.github.io

Overview			
	FF	RNN	CNN
- Basic computation - Typical data - Weights ties - Effective span	Matrix multiplies Feature vectors No Across input	Recurrent multiplications Sequences Sequential Long term	Convolutions Images Spatial Local receptive field
- How to train - Performance - Typical examples	Backprop No guarantee None	Backprop No guarantee Machine translation, speech tagging, text generation	Backprop No guarantee segmentation, image processing, object detection

October 24,
2017

Rob
Romijnders

Why

How

What

Data
Libraries
Start simple
Cloud training

Deploy

Wrap up

Questions

- Data
- Libraries
- Start simple
- Cloud training

Gather data

October 24,
2017

Rob
Romijnders

Why
How
What
Data
Libraries
Start simple
Cloud training

Deploy

Wrap up

Questions

1 SMALL:

- 1 MNIST: 60.000 samples
- 2 TIMIT: 630 speakers, 10 sentences

2 NORMAL:

- 1 Imagenet: 3.2 million images
- 2 FLICKR: 1 million images
- 3 Word vectors on all of Wikipedia

Semi supervised learning

October 24,
2017

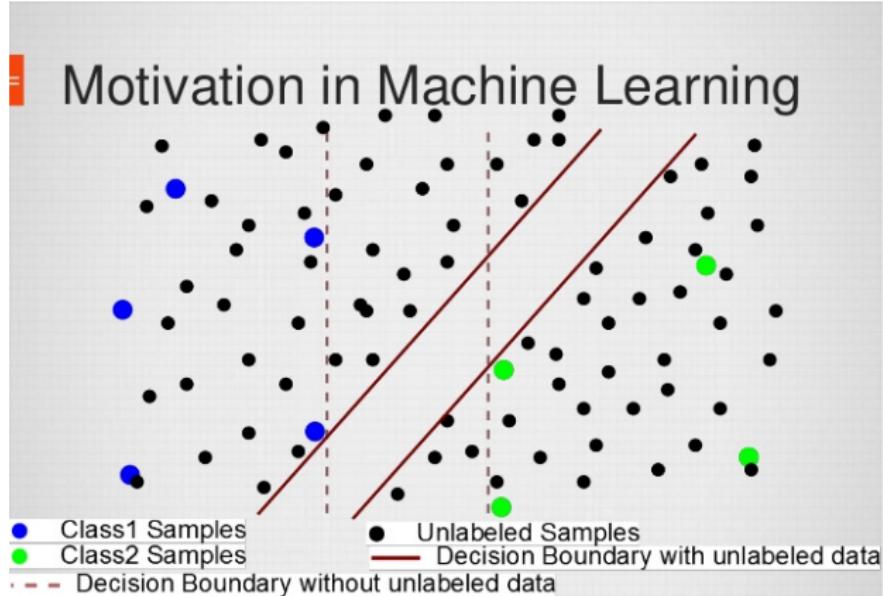
Rob
Romijnders

Why
How
What
Data
Libraries
Start simple
Cloud training

Deploy

Wrap up

Questions



:: Semi-Supervised Learning :: Lukas Tencer :: MTL Data ::



Figure: Credits: Lukas Tencer

MNIST

October 24,
2017

Rob
Romijnders

Why
How
What
Data
Libraries
Start simple
Cloud training

Deploy
Wrap up
Questions

Random Sampling of MNIST

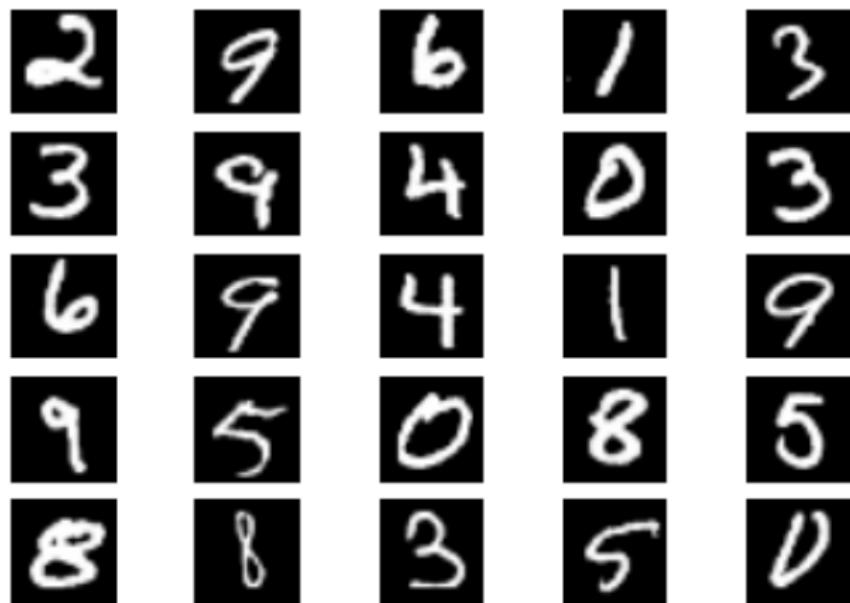


Figure: MNIST images

Code

October 24,
2017

Rob
Romijnders

Why

How

What

Data

Libraries

Start simple

Cloud training

Deploy

Wrap up

Questions

1 Low level

- 1 Tensorflow
- 2 Torch, pyTorch

2 High level

- 1 Keras
- 2 Scikit flow, pretty tensor, tfslim, ...

Train

October 24,
2017

Rob
Romijnders

Why

How

What

Data
Libraries

Start simple
Cloud training

Deploy

Wrap up

Questions

- 1 No bells and whistles**
- 2 Start simple**

Choose algorithm

October 24,
2017

Rob
Romijnders

Why

How

What

Data

Libraries

Start simple

Cloud training

Deploy

Wrap up

Questions

1 Convolution

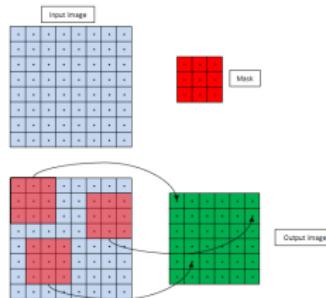


Figure: developer.amd.com

2 Recursion

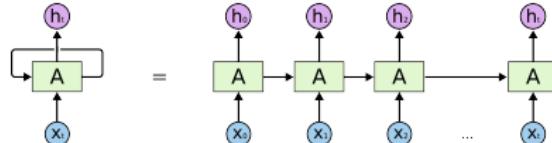


Figure: credits: colah.github.io

Bells and whistles

October 24,
2017

Rob
Romijnders

Why

How

What

Data

Libraries

Start simple

Cloud training

Deploy

Wrap up

Questions

- 1 Batch normalization, layer normalization
- 2 Dropout
- 3 Activation functions
- 4 Residual connections
- 5 etcetera, etcetera

Activation functions

October 24,
2017

Rob
Romijnders

Why

How

What

Data
Libraries

Start simple

Cloud training

Deploy

Wrap up

Questions

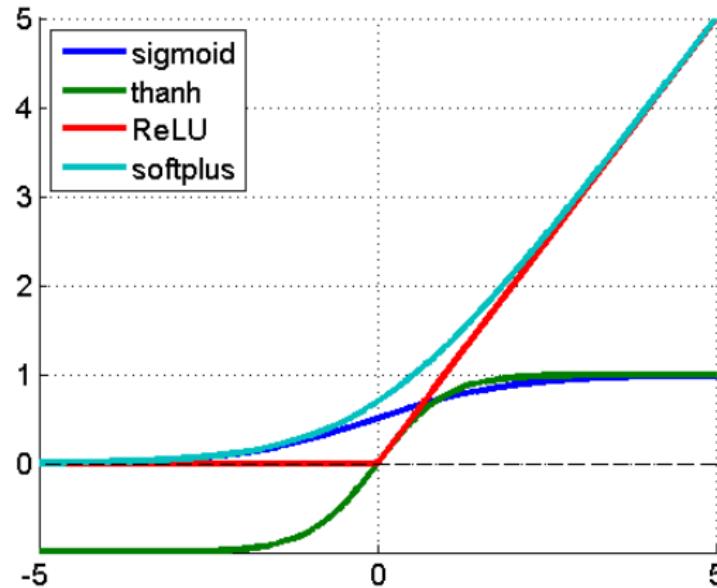


Figure: Common activation functions (credits Vanessa Imiloa)

Play with activation functions here

Cloud services

October 24,
2017

Rob
Romijnders

Why

How

What

Data

Libraries

Start simple

Cloud training

Deploy

Wrap up

Questions

- 1 AWS**
- 2 Azure**
- 3 Google cloud**

Docker

October 24,
2017

Rob
Romijnders

Why

How

What

Data

Libraries

Start simple

Cloud training

Deploy

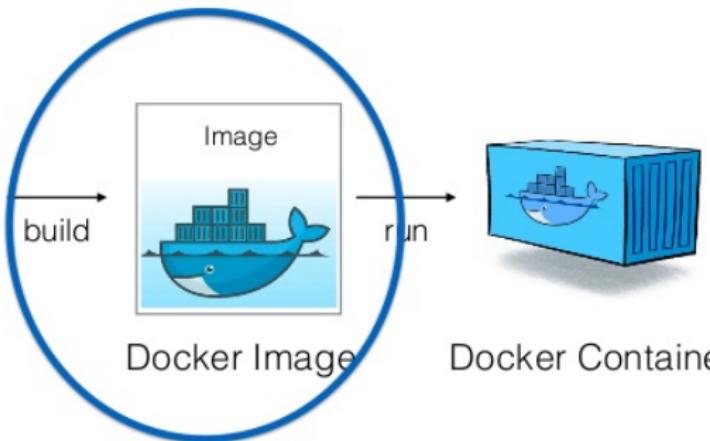
Wrap up

Questions

Single Container Docker Workflow

```
FROM ubuntu:16.04
MAINTAINER Rob Romijnders <rob@romijnders.net>
ADD . /app
WORKDIR /app
RUN apt-get update
RUN apt-get install -y curl
RUN curl -s https://raw.githubusercontent.com/rob-romijnders/docker-training/master/app/app.py > app.py
CMD ["python", "app.py"]
```

Dockerfile



Docker pipeline

October 24,
2017

Rob
Romijnders

Why

How

What

Data

Libraries

Start simple

Cloud training

Deploy

Wrap up

Questions

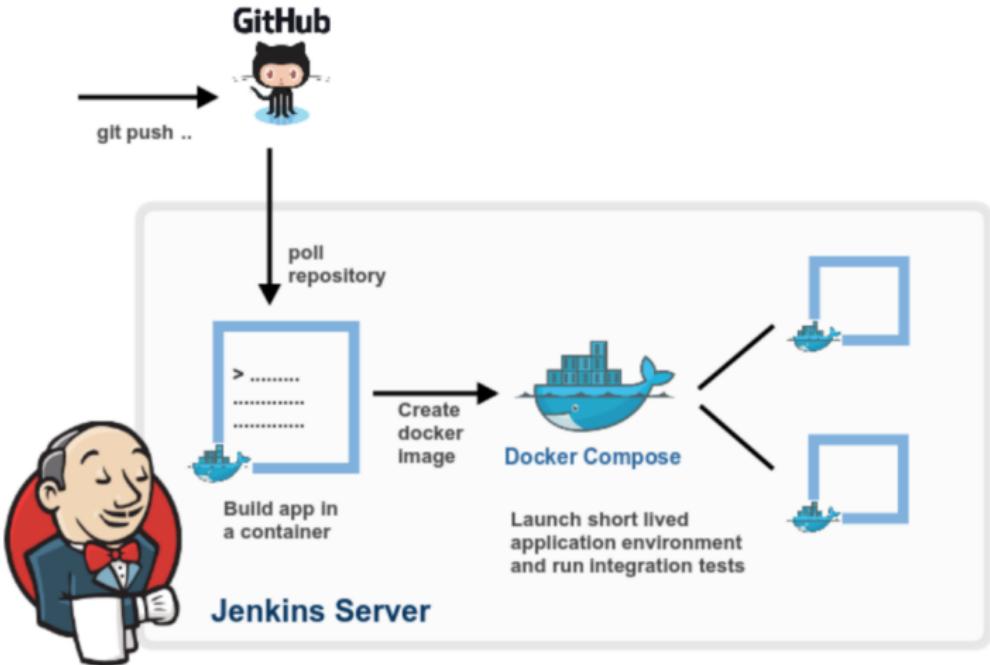


Figure: Source: rancher.com

Gradient Descent

October 24,
2017

Rob
Romijnders

Why
How
What
Data
Libraries
Start simple
Cloud training

Deploy

Wrap up

Questions

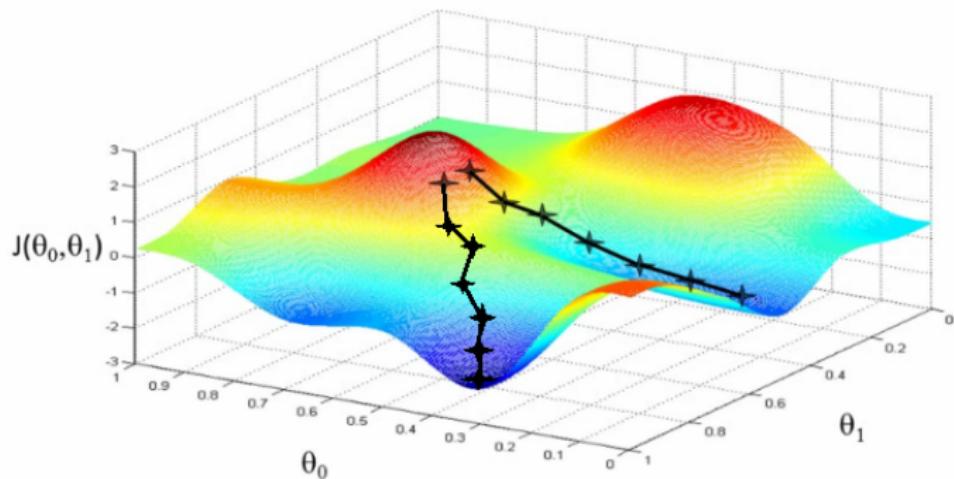


Figure: Credits: ML blog Vasilis Vryniotis

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Fast inference
Micro services
Post processing

Wrap up

Questions

Deploy

DIY

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Fast inference
Micro services
Post processing

Wrap up

Questions

- 1 Fast inference at testtime**
- 2 Microservices**
- 3 Post processing**

Quantization

October 24,
2017

Rob
Romijnders

Why
How
What
Deploy
Fast inference
Micro services
Post processing
Wrap up
Questions

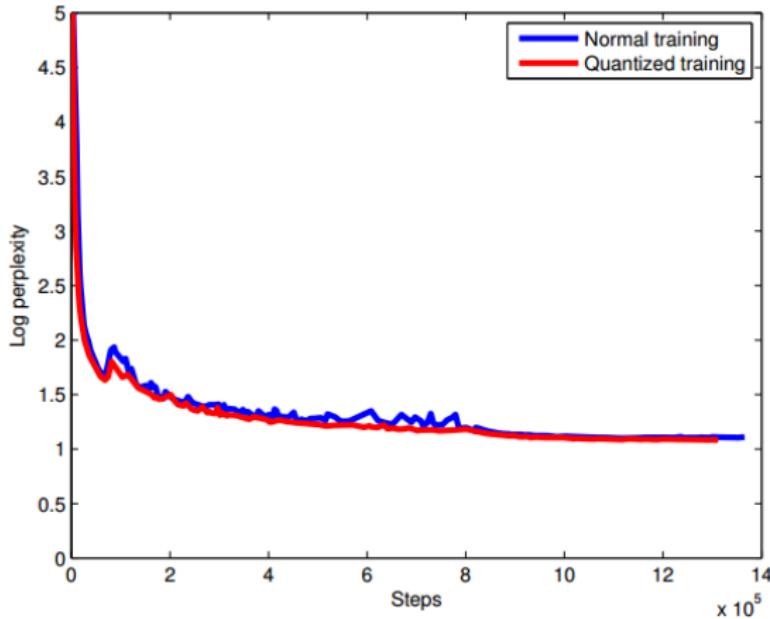


Figure: Credits: Arxiv GNMT paper

Quantization

October 24,
2017

Rob
Romijnders

Why
How
What
Deploy
Fast inference
Micro services
Post processing
Wrap up
Questions

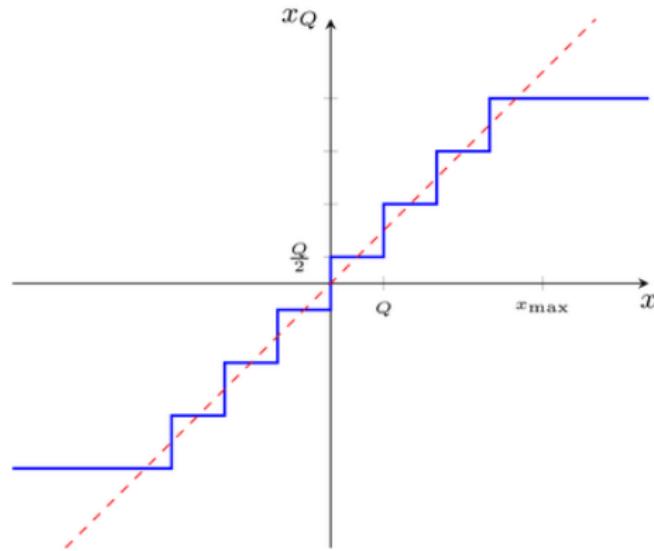


Figure: Credits: Sascha Spors

Speed up

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Fast inference

Micro services

Post processing

Wrap up

Questions

	BLEU	Log Perplexity	Decoding time (s)
CPU	31.20	1.4553	1322
GPU	31.20	1.4553	3028
TPU	31.21	1.4626	384

Figure: Credits: [Arxiv GNMT paper](#)

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Fast inference

Micro services

Post processing

Wrap up

Questions

Micro services

Micro services

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Fast inference

Micro services

Post processing

Wrap up

Questions

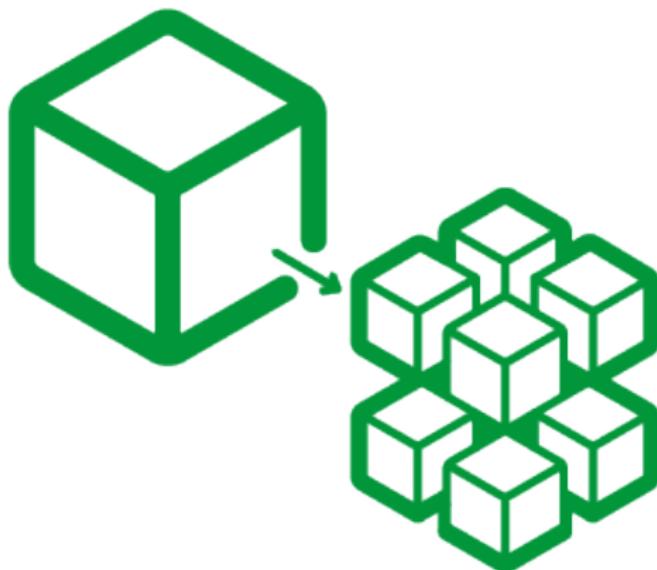


Figure: Source: nginx.com

Post processing

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Fast inference

Micro services

Post processing

Wrap up

Questions

- 1 Plot low confidence / high uncertainty**
- 2 Plot confusion matrix**
- 3 Inspect neurons**

Regularization

October 24,
2017

Rob
Romijnders

Why
How
What
Deploy
Fast inference
Micro services
Post processing
Wrap up
Questions



Figure: Example for regularization

October 24,
2017

Rob
Romijnders

Why
How
What
Deploy
Fast inference
Micro services
Post processing
Wrap up
Questions



October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

Wrap up

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

Questions

Further learning

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

1 Blogs

- 1 [The Unreasonable Effectiveness of Recurrent Neural Networks, Andrej Karpathy](#)
- 2 [Chris Olah's blog colah.github.io](#)

2 Courses

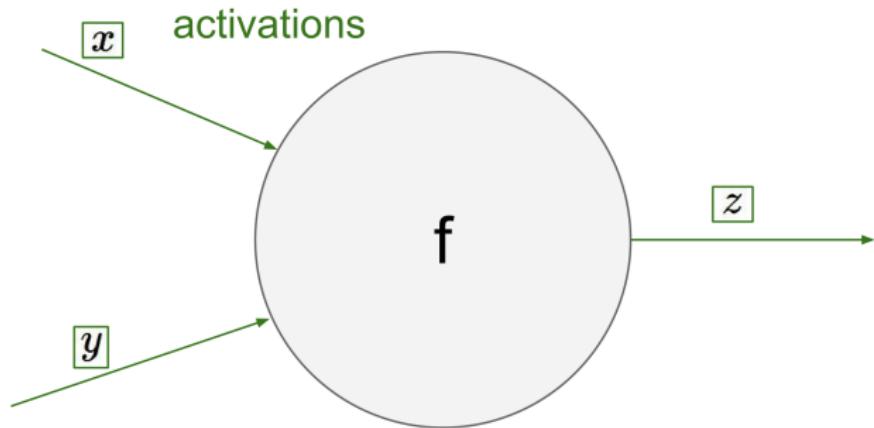
- 1 [Course on CNN: cs231n](#)
- 2 [Udacity on Deep learning](#)
- 3 [Fast.ai: Deep learning for Coders](#)

3 My page (robromijnders.github.io)

October 24,
2017

Rob
Romijnders

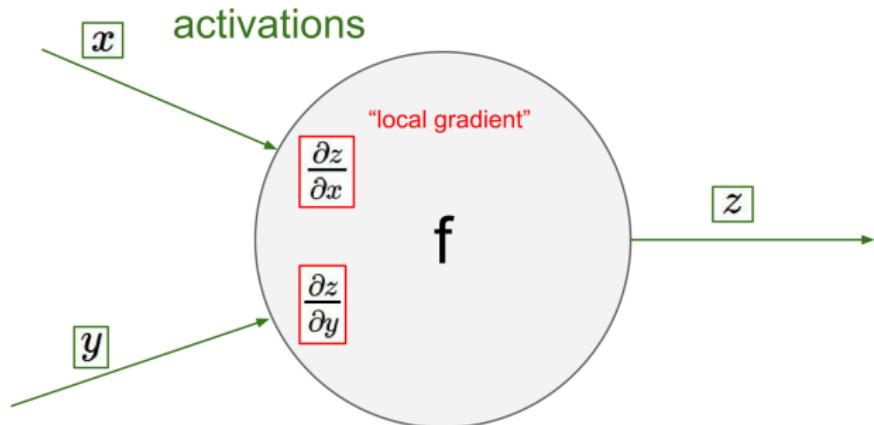
Why
How
What
Deploy
Wrap up
Questions



October 24,
2017

Rob
Romijnders

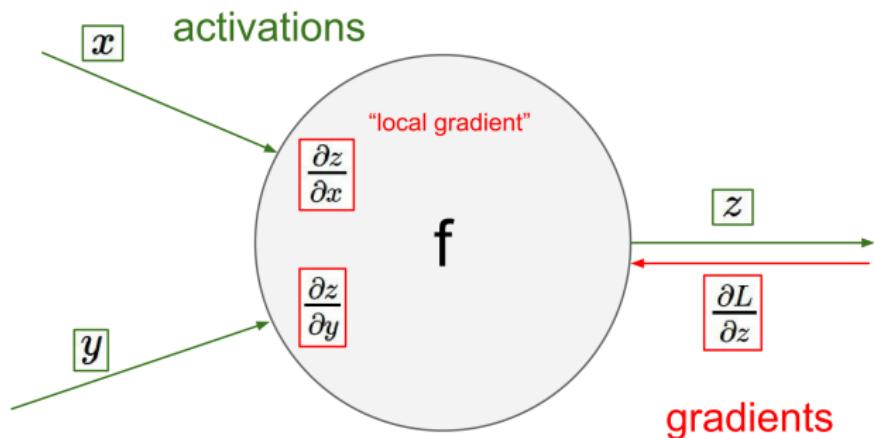
Why
How
What
Deploy
Wrap up
Questions

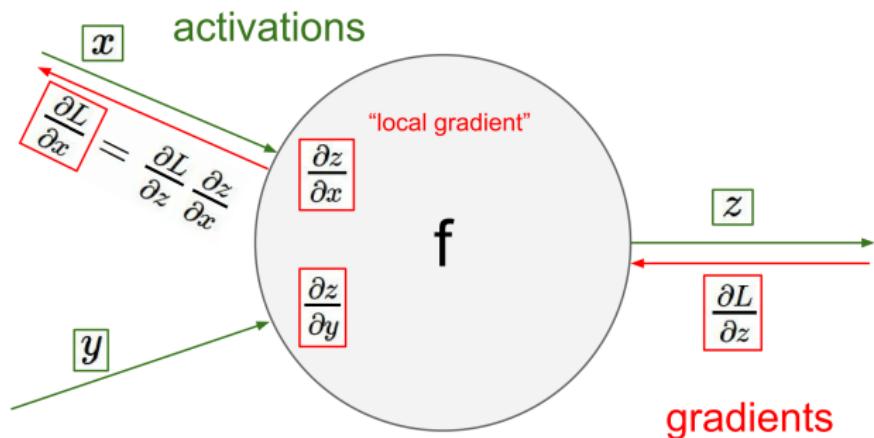


October 24,
2017

Rob
Romijnders

Why
How
What
Deploy
Wrap up
Questions

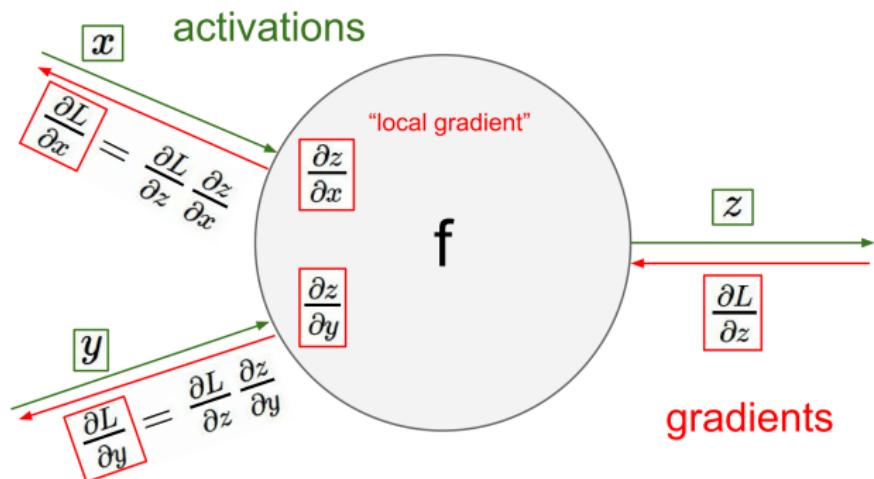


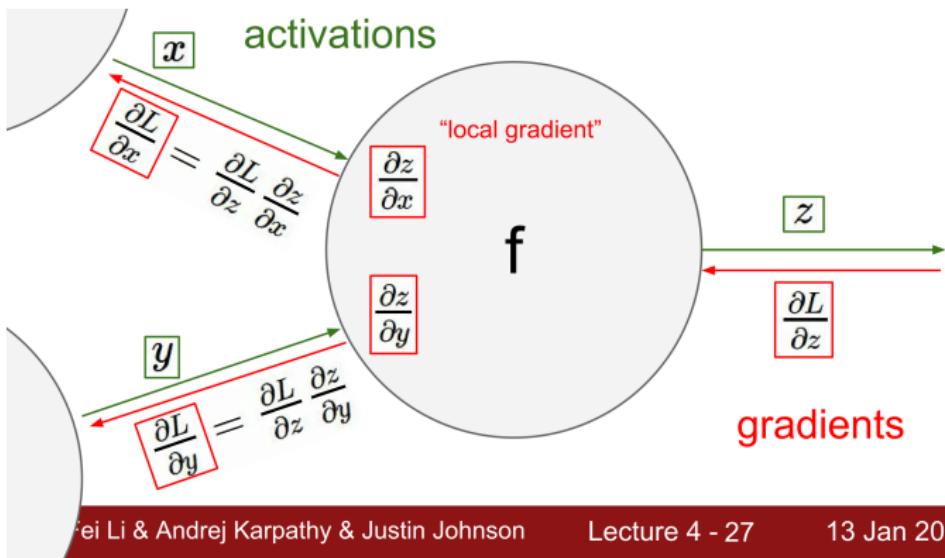


October 24,
2017

Rob
Romijnders

Why
How
What
Deploy
Wrap up
Questions





Softmax

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

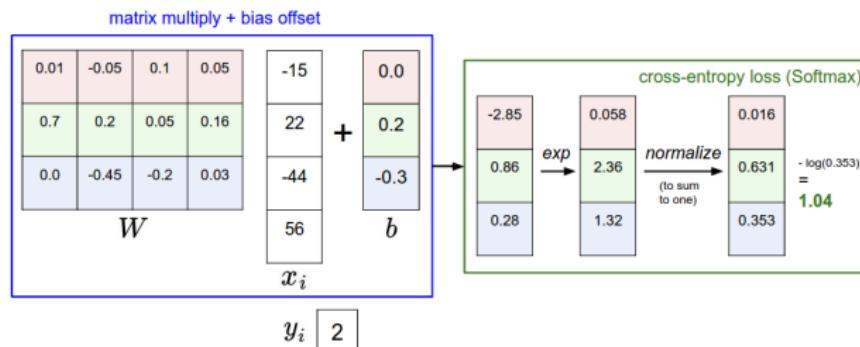


Figure: Adapted from cs231n.github.io/linear-classify/

Example

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

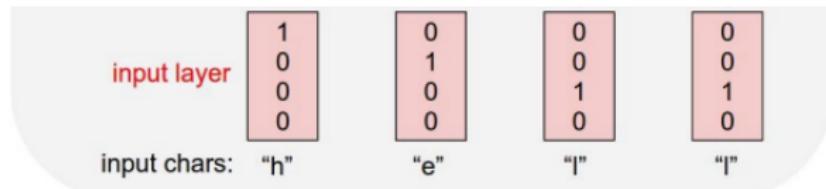


Figure: cs231n, Andrej Karpathy

Example

October 24,
2017

Rob
Romijnders

Why

How

What

Deploy

Wrap up

Questions

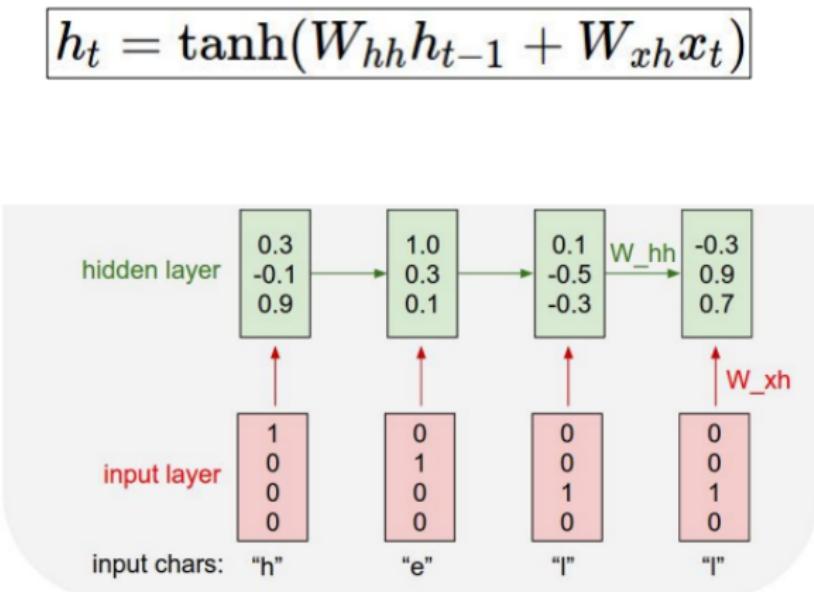


Figure: cs231n, Andrej Karpathy

Example

October 24,
2017

Rob
Romijnders

Why
How
What
Deploy
Wrap up
Questions

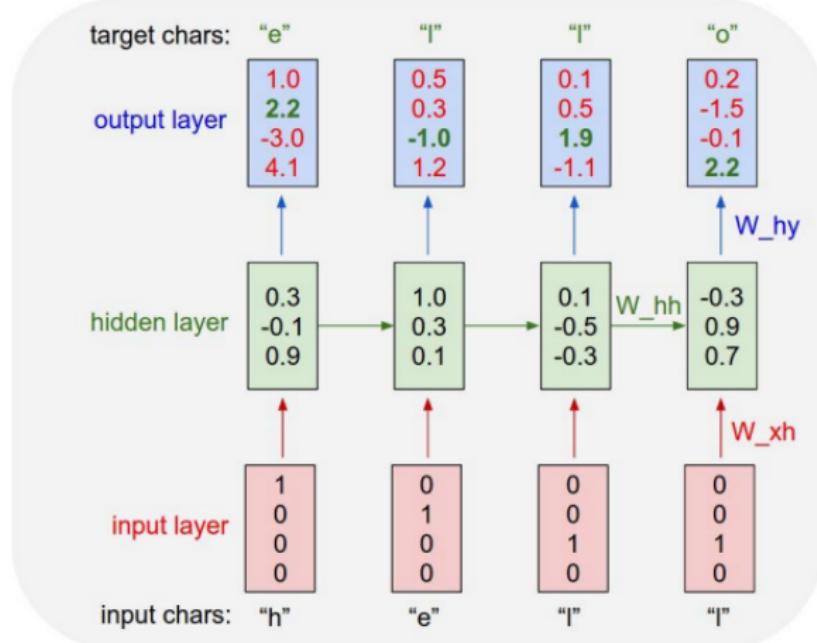


Figure: cs231n, Andrej Karpathy