Racket for Deep Learning

Charles Earl
Data Scientist, Automattic
charles.earl@automattic.com

AUTOMATTIC
We are passionate about making the web a better place.

Contribute at https://github.com/charlescearl/DeepRacket

AUTOMATTIC

We are passionate about making the web a better place.

Outline

- Overview of deep learning
- Potential of Racket for deep learning
- Progress on the DeepRacket library
- What should the next steps be?

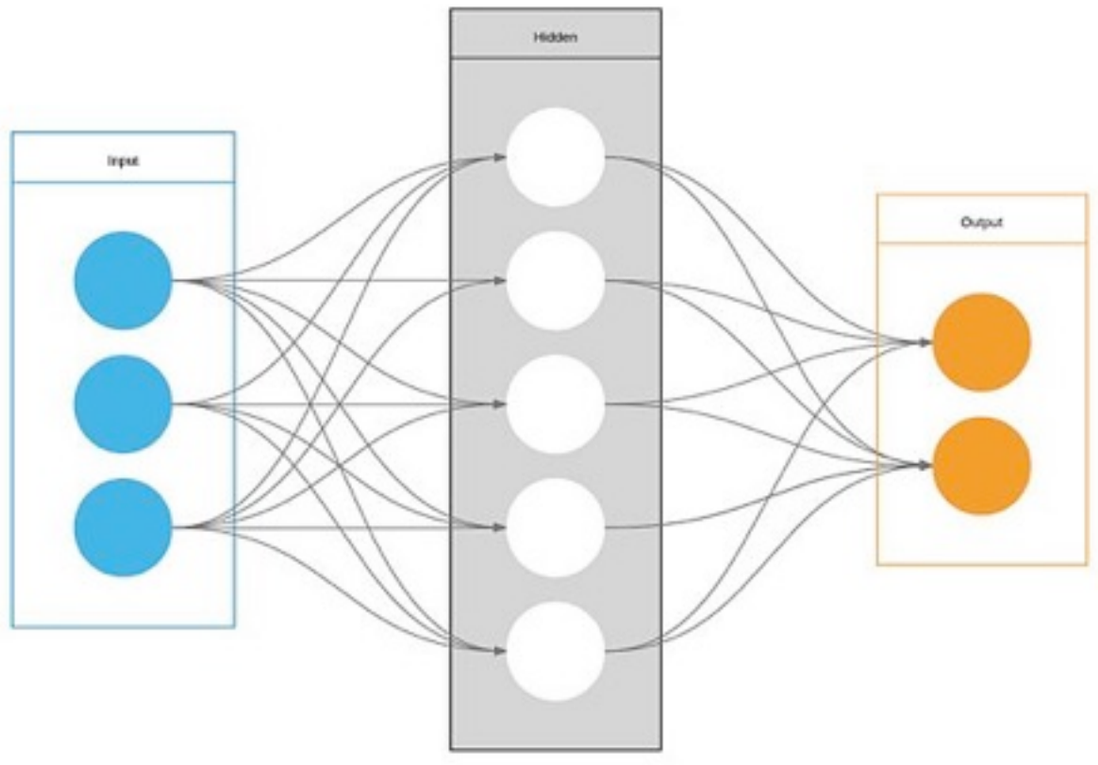
AUTOMATTIC
We are passionate about making the web a better place.

Overview of Deep Learning

- What are neural networks?
- What is deep learning?
- How does deep learning work?

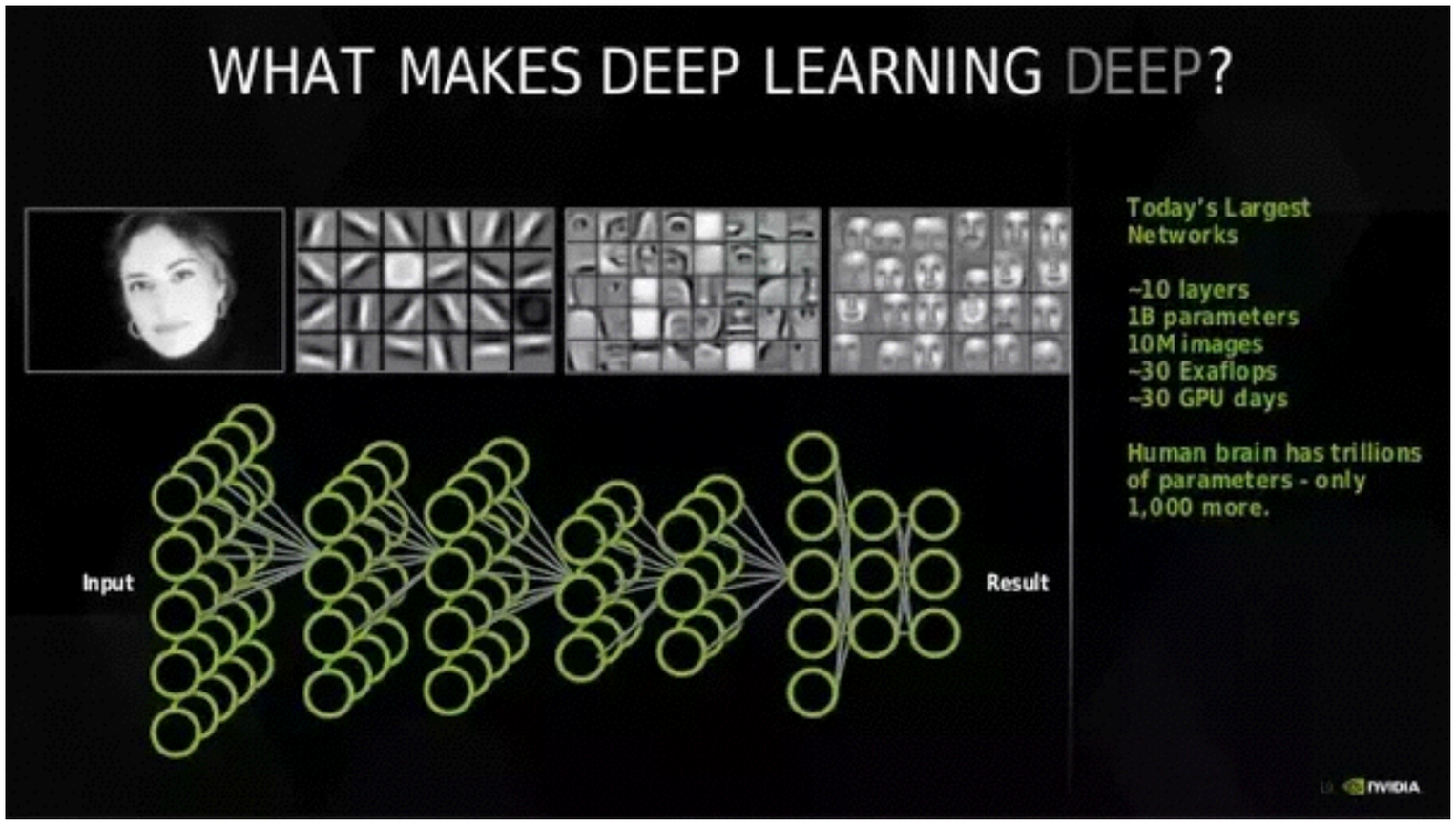
AUTOMATTIC
We are passionate about making the web a better place.

What are neural networks?



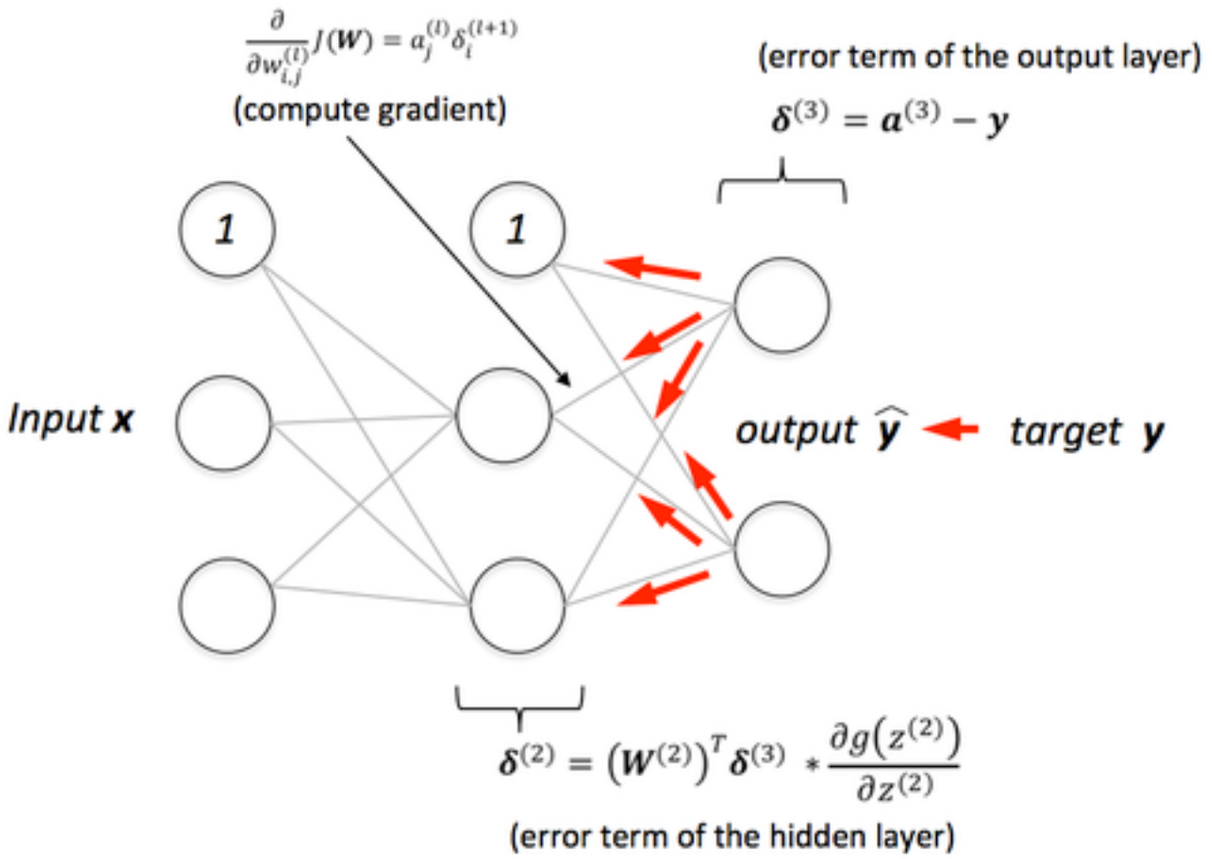
AUTOMATTIC
We are passionate about making the web a better place.

What is deep learning?



From Nvidia

How does deep learning work?



AUTOMATTIC
We are passionate about making the web a better place.

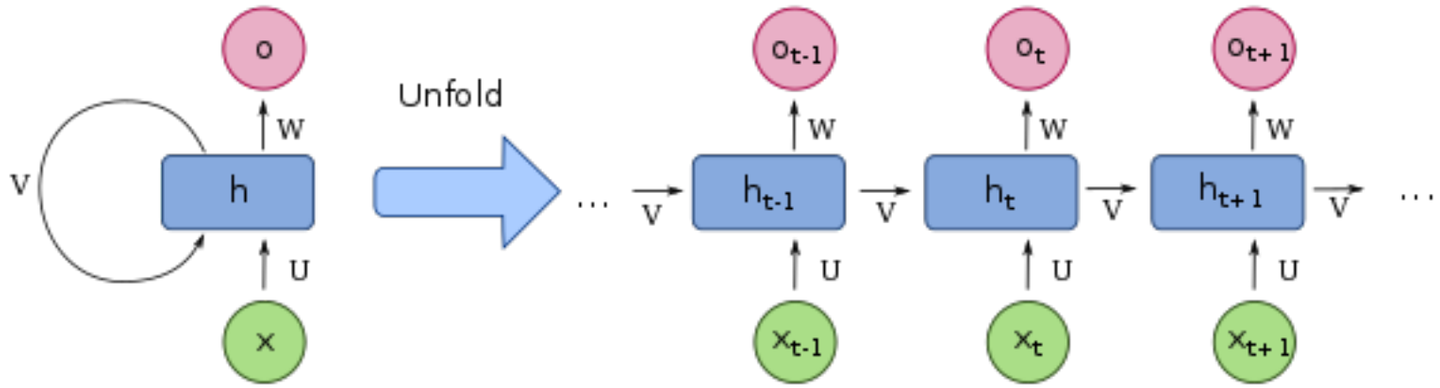
from https://sebastianraschka.com/faq/docs/visual-backpropagation.html

Potential of Racket for Deep Learning

- How does functional programming apply to deep leaning?
- Specifying deep networks dynamically.
- Specifying and running in Racket

AUTOMATTIC
We are passionate about making the web a better place.

Recurrent Nets



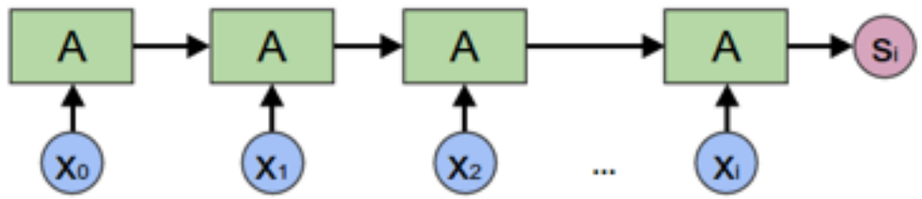
$$h_t = W \cdot h_{t-1}$$

$$o_t = tanh(h_t + U \cdot x_t)$$

AUTOMATTIC
We are passionate about making the web a better place.

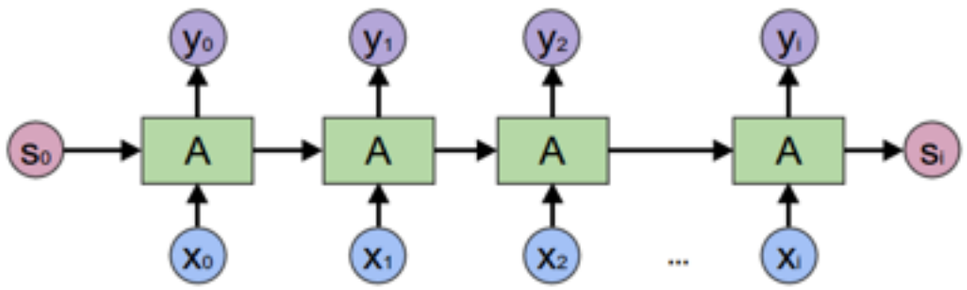
Thinking of Deep Learning Functionally (Colah, 2015 & Balduzzi 2016)

Encoding RNN



foldl A s

RNN

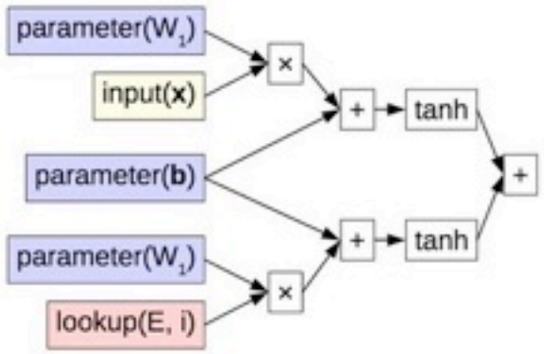


mapAccumR a s

AUTOMATTIC
We are passionate about making the web a better place.

from Neural Networks, Types, and Functional Programming
http://colah.github.io/posts/2015-09-NN-Types-FP/

As dynamic graph computation



(a) Computation graph

```
1 b = dy.parameter(cg, b_param)
2 W1 = dy.parameter(cg, W1_param)
3 W2 = dy.parameter(cg, W2_param)
4 x = dy.vectorInput(x_vec)
5 e = e_param[i]
6 g = dy.tanh(W1 * x + b) + dy.tanh(W2 * e + b)
7 g_val = g.value()
8 g.backward()
```

(b) Corresponding code

Figure 4: An example of a computation graph for $g(\boldsymbol{x}, j) = \tanh(W_1 * \boldsymbol{x} + \boldsymbol{b}) + \tanh(W_2 * e_j + \boldsymbol{b})$, and the corresponding code.

from Dynet, https://github.com/clab/dynet

Specifying and running in Racket

Specify the network

```
(let ([W  (def-filter (3 2))]
      [ x    (def-tensor (2))]
      …
      [ sgd  (create-stock-grad-optimizer )]
[ rnn  (def-expr (+ (* W x) (* U h) b))  ])
```

Now train it!

```
(for ([i (range epochs)])
     (compute-fwd rnn)
     (compute-backward rnn (loss yhat y))
     (compute-backward rnn)
     (update-weights sgd rnn)))
```

AUTOMATTIC

We are passionate about making the web a better place.

Benefits

- Giving students exposure to deep learning
- Exploration of new functional architectures
- New paradigms for machine learning (e.g. alternatives to backprop)
- CUDA/OpenCL access for Racket

Progress on the DeepRacket library

AUTOMATTIC

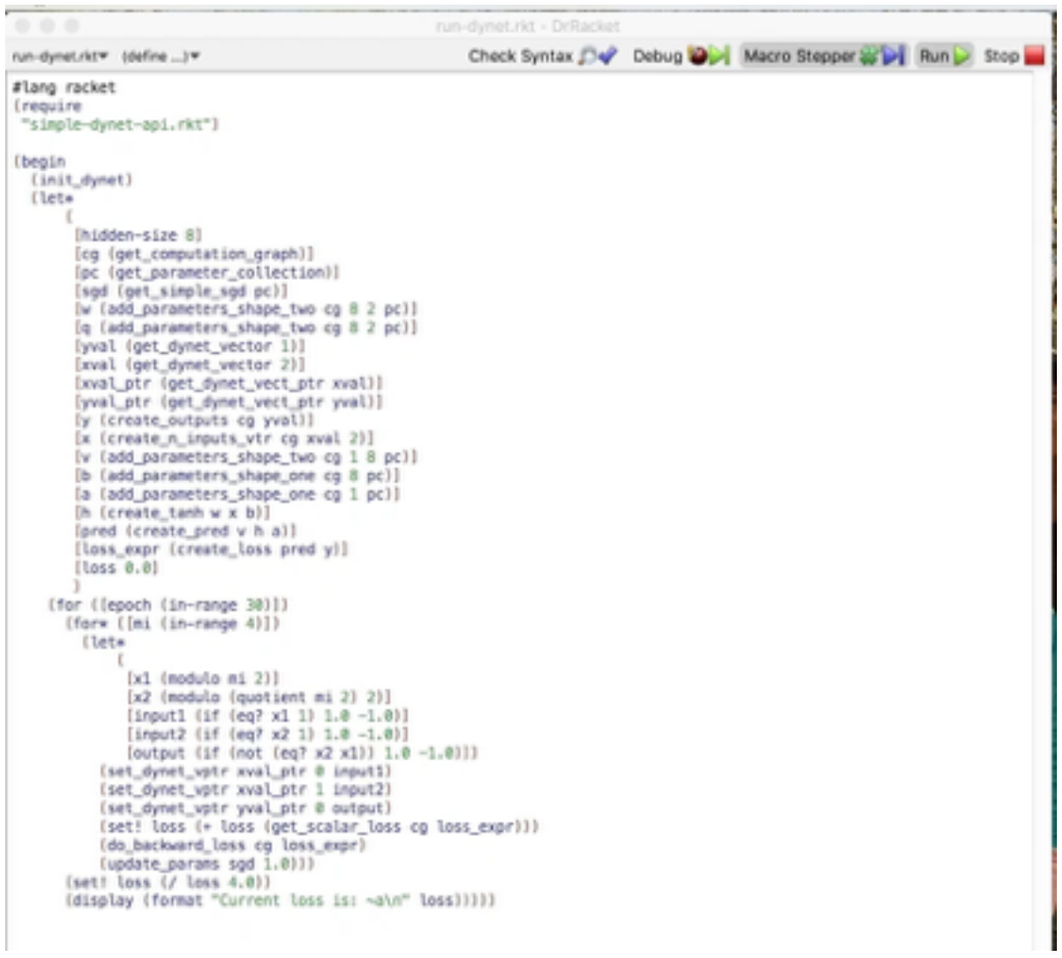
We are passionate about making the web a better place.

DeepRacket

- A cudnn wrapper
  - Built with typed Racket, Math and FFI libraries
  - Specification of RNNs and forward (estimator) and backward (gradient) computation
  - Still need loss and optimization (SGD)!
- A dynet wrapper
  - Specify selected networks
  - Training

AUTOMATTIC
We are passionate about making the web a better place.

A quick DeepRacket demo



AUTOMATTIC

We are passionate about making *the web* a better place.

What’s next?

AUTOMATTIC

We are passionate about making the web a better place.

Useful directions?

- Better Syntax!
- Device agnostic
  - GPU and CPU
  - OpenCL
- Dynamic graph computation

https://github.com/charlescearl/DeepRacket

AUTOMATTIC

We are passionate about making the web a better place.