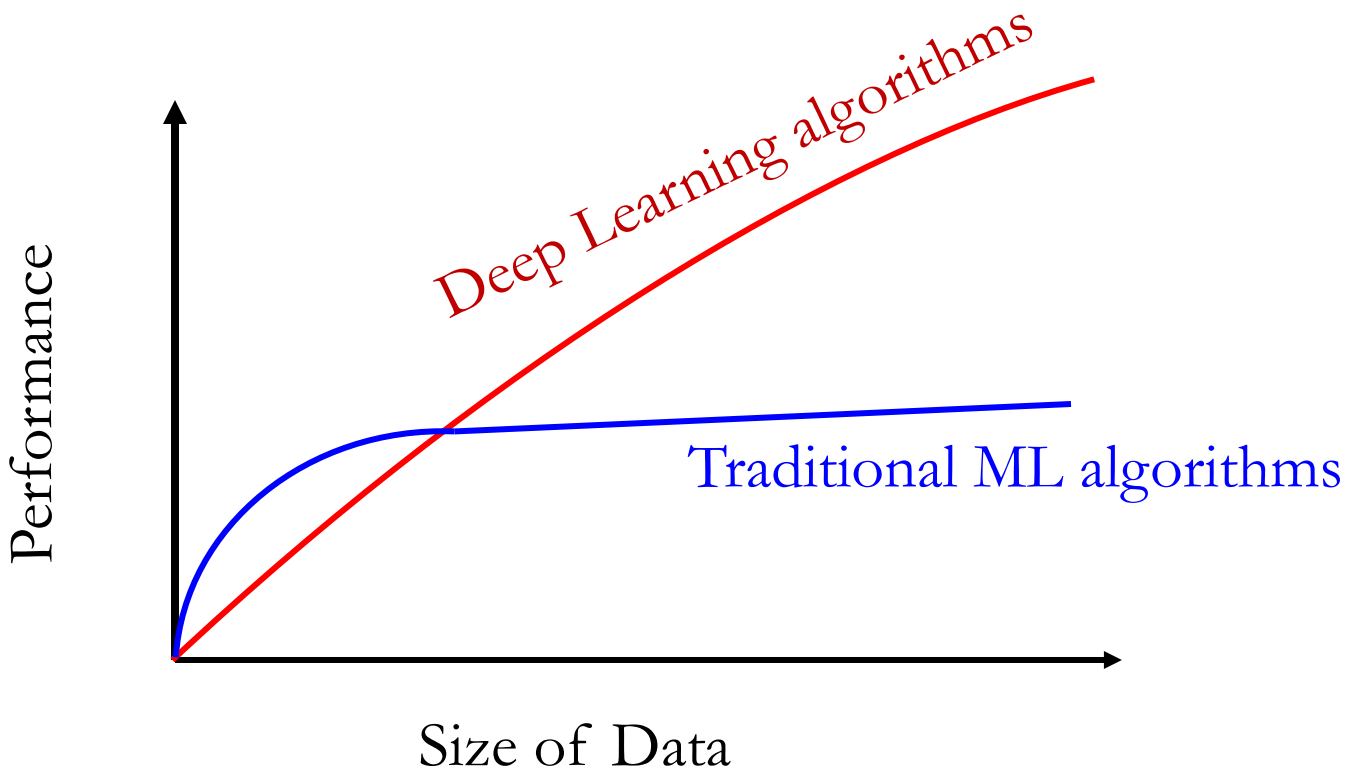
Deep Learning
Lecture-9&10

Asst. Lect. Ali Al-khawaja

2025-2026

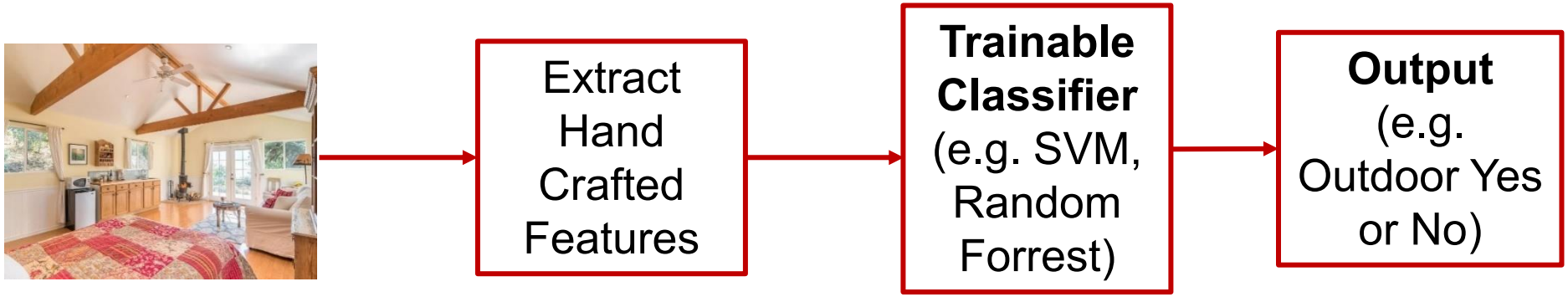
Class Room

Performance vs Sample Size



Supervised Learning

- Traditional pattern recognition models work with hand crafted features and relatively simple trainable classifiers.



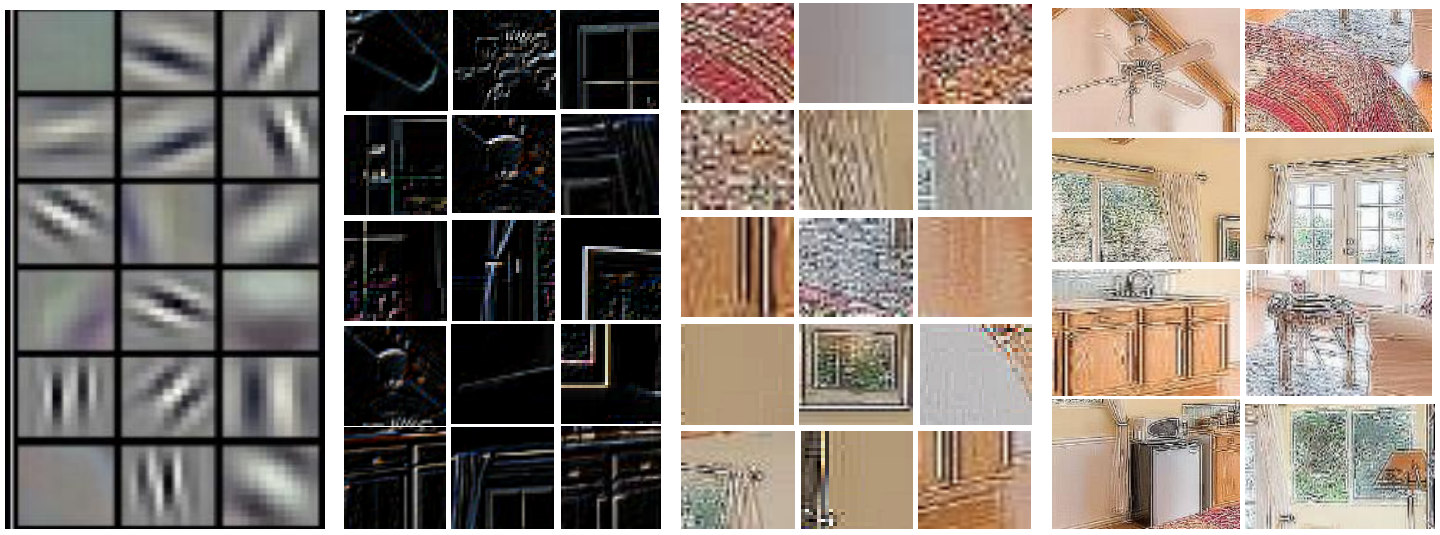
Limitations

- Very tedious and costly to develop hand crafted features.
- The hand-crafted features are usually highly dependents on one application.

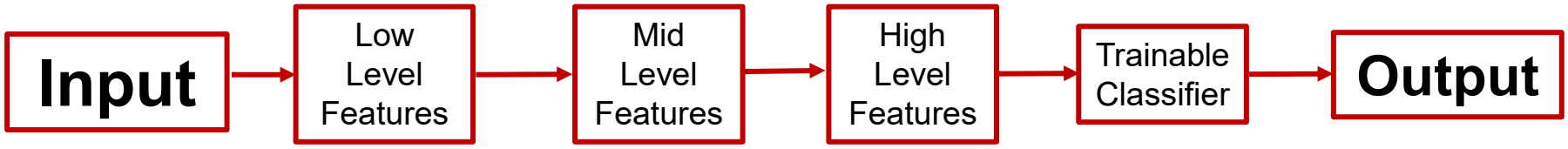
Deep Learning

- Deep learning has an **inbuilt automatic multi stage feature learning process** that learns rich hierarchical representations (i.e. features).

Low Level Features → Mid Level Features → High Level Features → Trainable Classifier → **Output (e.g. outdoor, indoor)**



Deep Learning



- Image

  Pixel → Edge → Texture → Motif → Part → Object

- Text

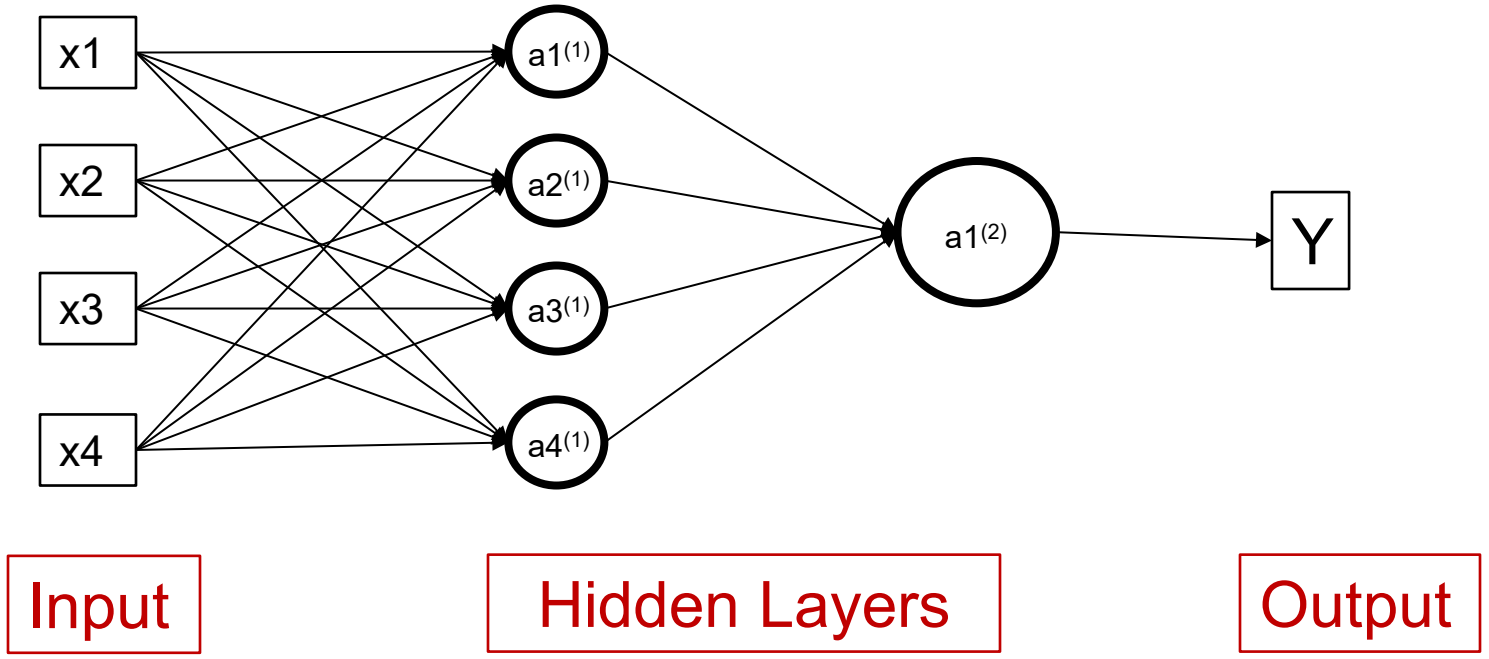
  Character → Word → Word-group → Clause → Sentence → Story

- Each module in Deep Learning transforms its input representation into a higher-level one, in a way similar to human cortex.

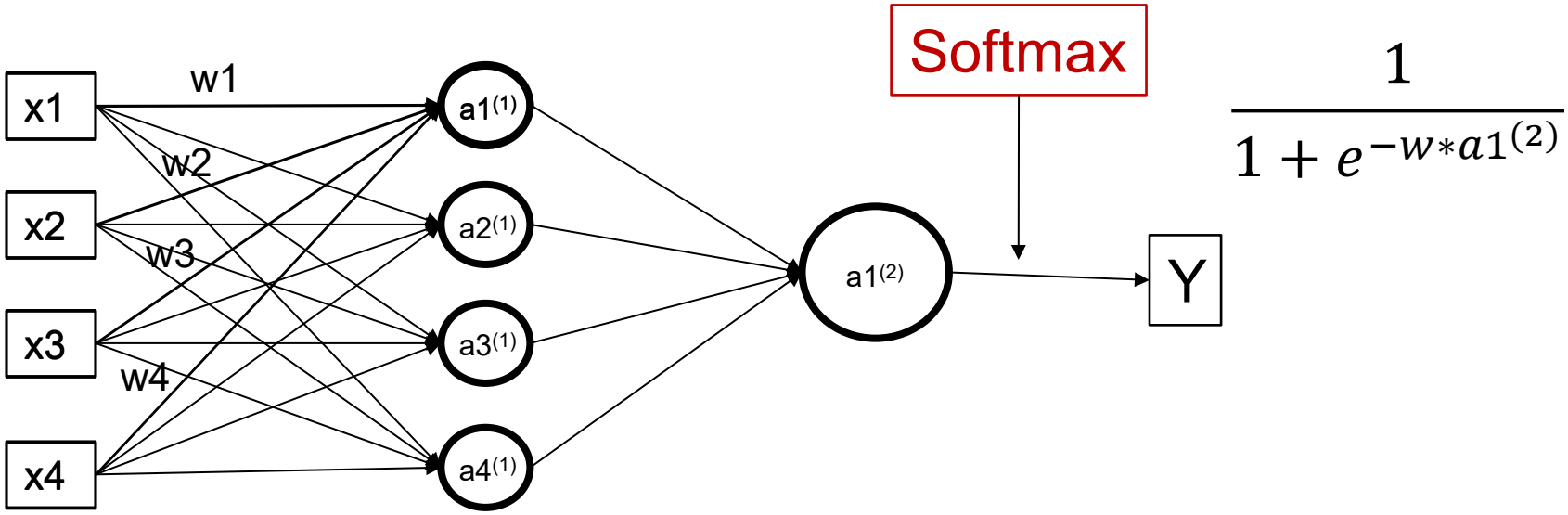
Let us see how it all works!

A Simple Neural Network

- An Artificial Neural Network is an information processing paradigm that is inspired by the biological nervous systems, such as the human brain's information processing mechanism.



A Simple Neural Network



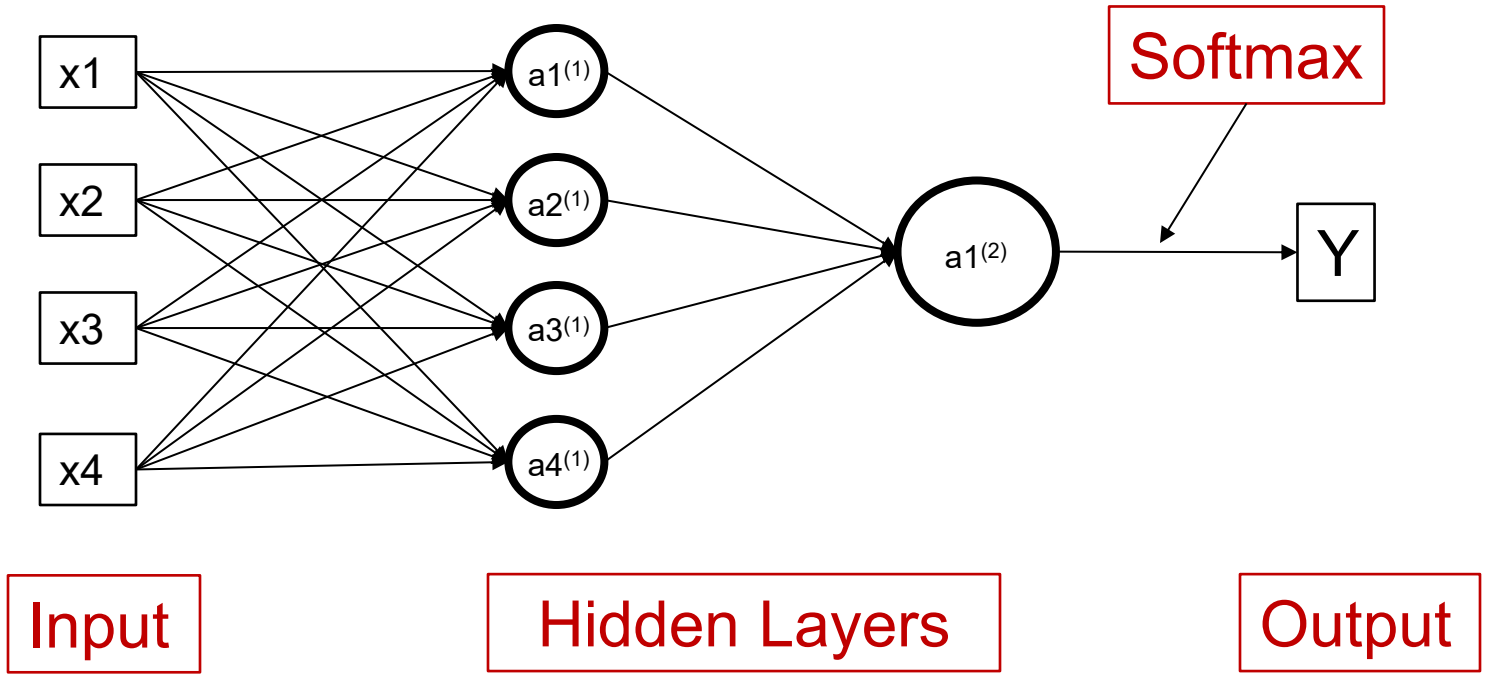
$$a1^{(1)} = f(w1 * x1 + w2 * x2 + w3 * x3 + w4 * x4)$$

f( ) is activation function: Relu or sigmoid

$$Relu: \max(0, x)$$

$$a1^{(1)} = max(0, w1 * x1 + w2 * x2 + w3 * x3 + w4 * x4)$$

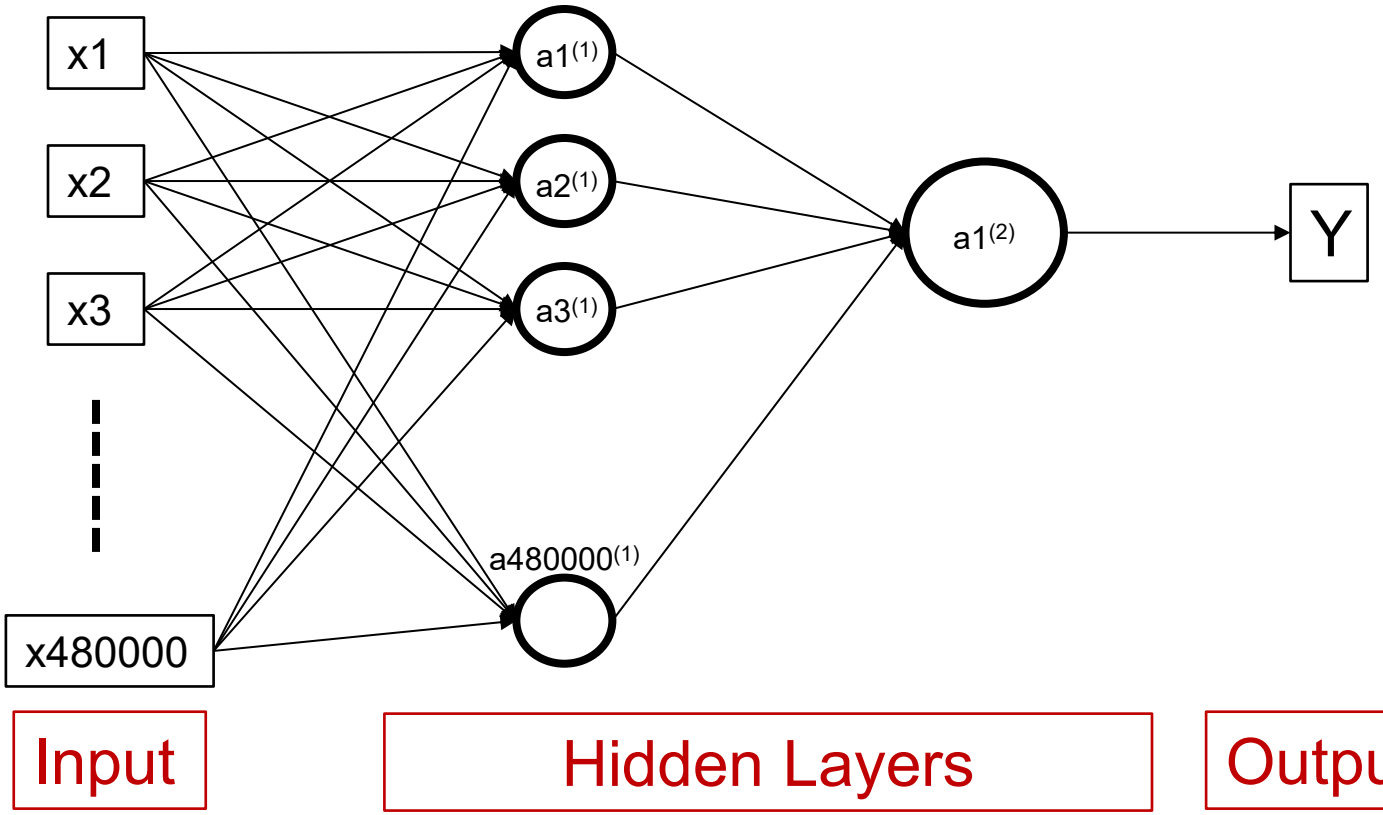
Number of Parameters



$$4*4 + 4 + 1$$

If the input is an Image?

400 X 400 X 3



Number of Parameters

480000*480000 + 480000 +1 = **approximately 230 Billion !!!**

480000*1000 + 1000 +1 = approximately **480 million !!!**

Let us see how convolutional layers help.

Convolutional Layers

- Filter  $\begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

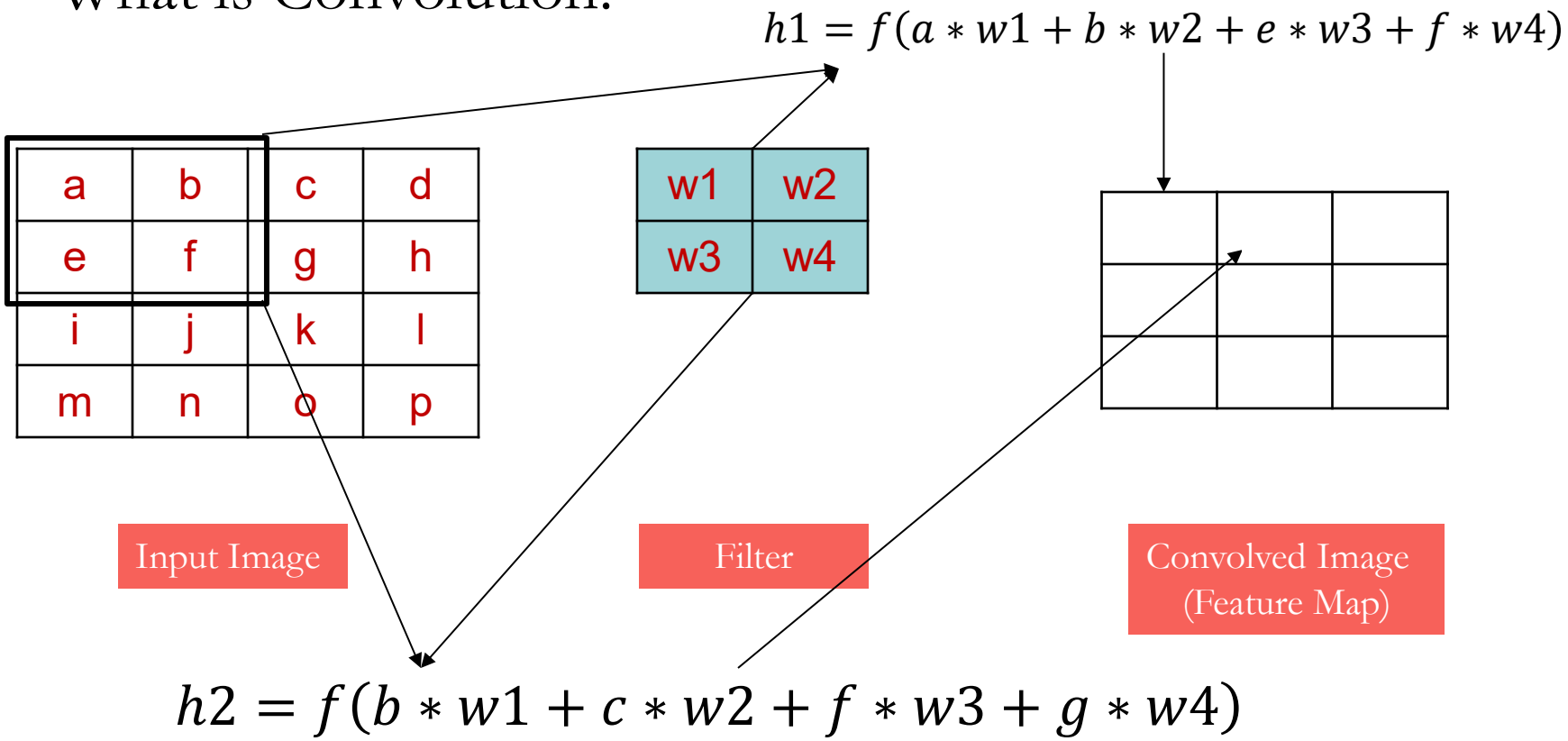
Input Image

Convoluted Image

- Inspired by the neurophysiological experiments conducted by Hubel and Wiesel 1962.

Convolutional Layers

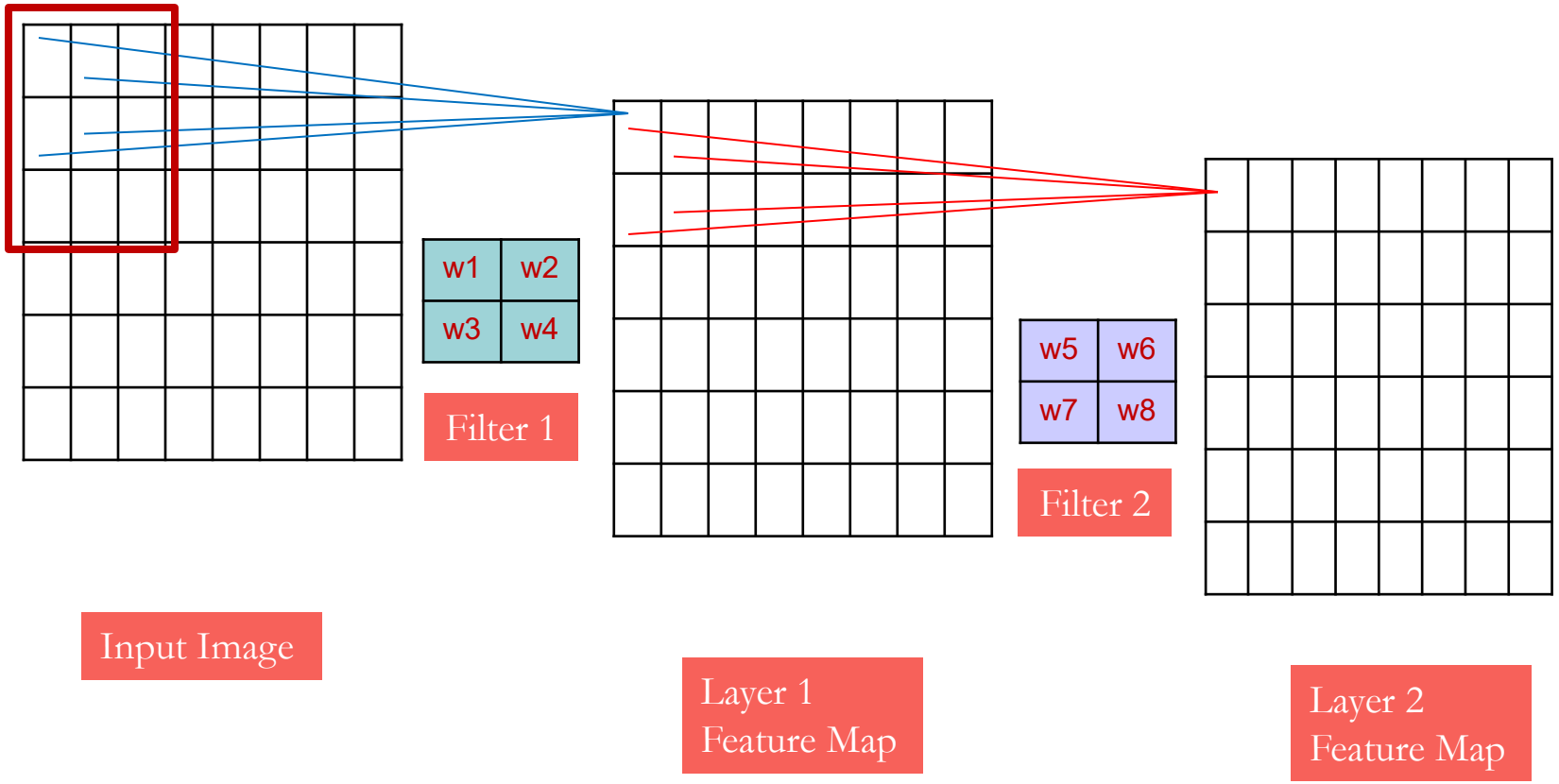
- What is Convolution?



Number of Parameters for one feature map = 4

Number of Parameters for 100 feature map = 4*100

Lower Level to More Complex Features



- In Convolutional neural networks, hidden units are only connected to local receptive field.

Pooling

- Max pooling: reports the maximum output within a rectangular neighborhood.
- Average pooling: reports the average output of a rectangular neighborhood.

1	3	5	3
4	2	3	1
3	1	1	3
0	1	0	4

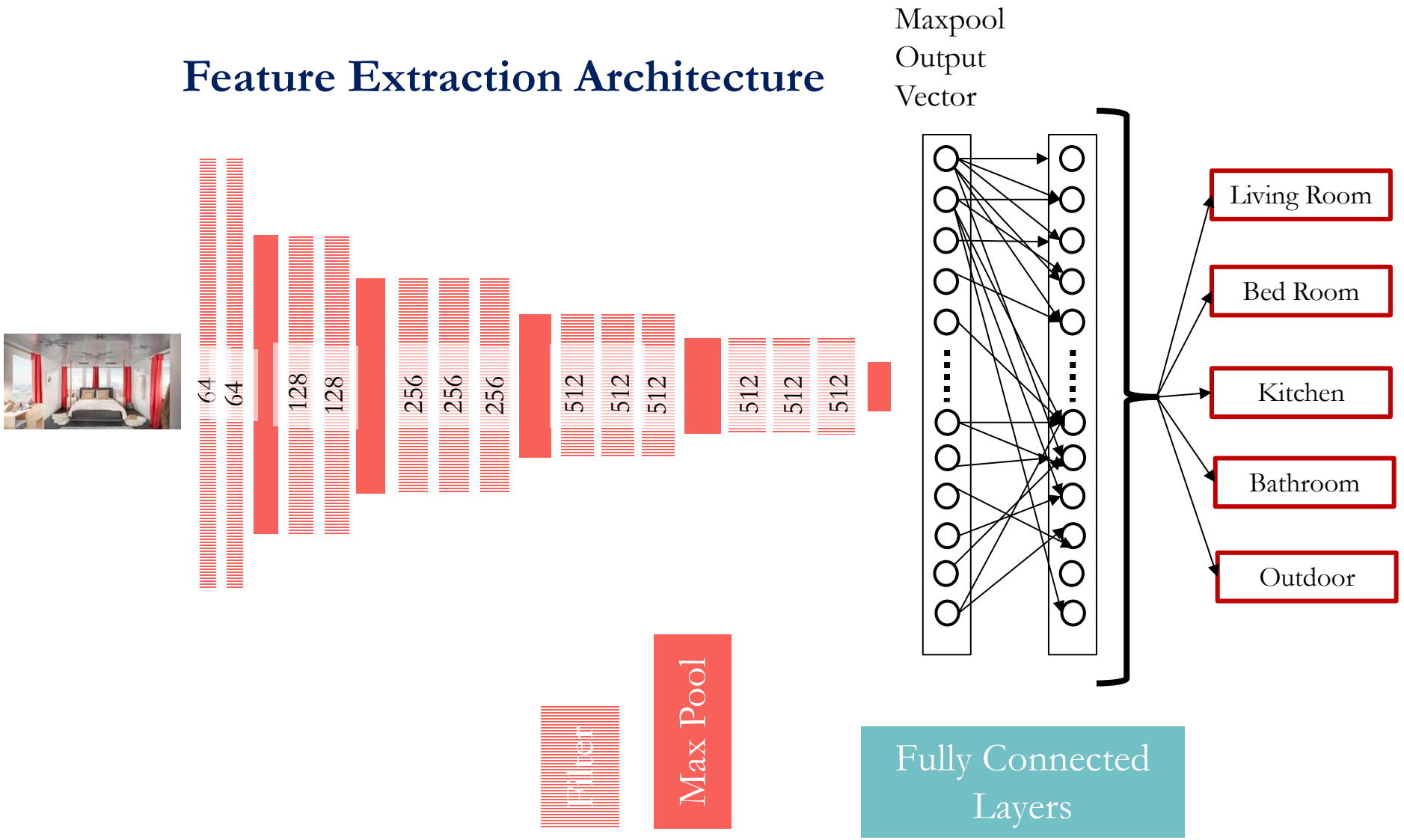
Input Matrix

MaxPool with 2X2 filter with stride of 2

4	5
3	4

Output Matrix

Convolutional Neural Network

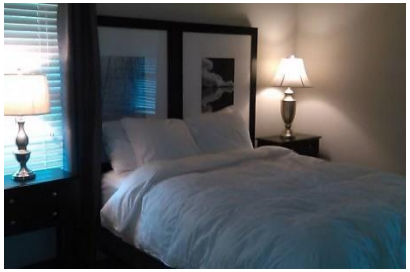
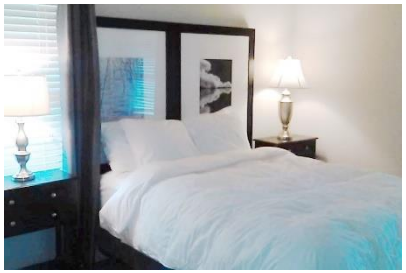


Convolutional Neural Networks

- **Output:** Binary, Multinomial, Continuous, Count
- **Input:** fixed size, can use padding to make all images same size.
- **Architecture:** Choice is ad hoc
  - requires experimentation.
- **Optimization:** Backward propagation
  - hyper parameters for very deep model can be estimated properly only if you have billions of images.
    - Use an architecture and trained hyper parameters from other papers (Imagenet or Microsoft/Google APIs etc)

Automatic Colorization of Black and White Images

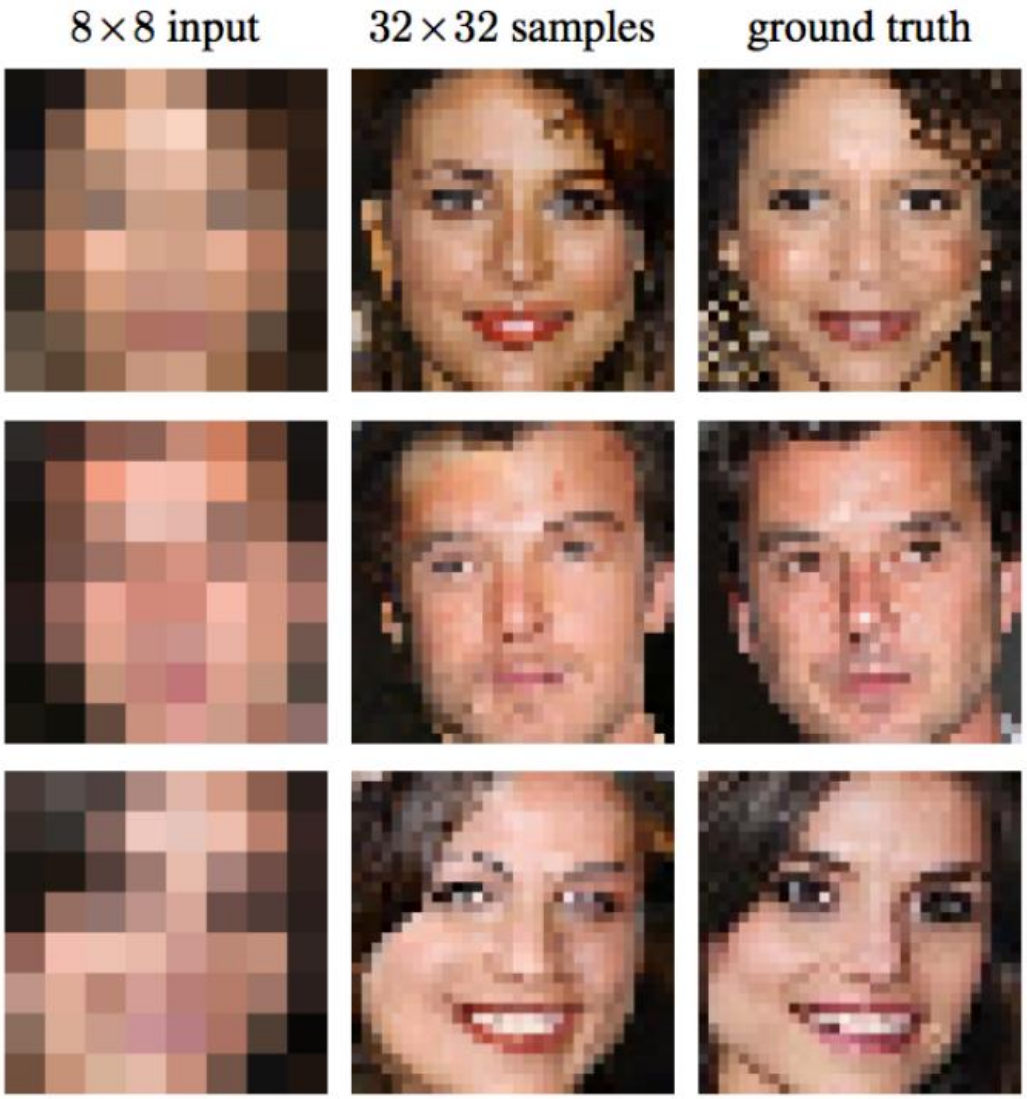
Optimizing Images

Post Processing Feature Optimization
(Color Curves and Details)


Post Processing Feature Optimization
(Illumination)


Post Processing Feature Optimization
(Color Tone: Warmness)



8x8 pixel photos were inputted into a Deep Learning network which tried to guess what the original face looked like. As you can see it was fairly close (the correct answer is under "ground truth").

Recurrent Neural Networks

Why RNN?

The limitations of the Convolutional Neural Networks

- Take fixed length vectors as input and produce fixed length vectors as output.
- Allow fixed amount of computational steps.

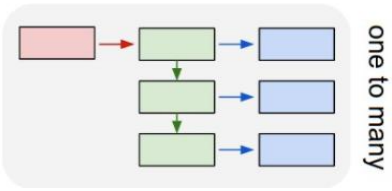
We need to model the data with temporal or sequential structures and varying length of inputs and outputs

e.g.

This movie is ridiculously good.

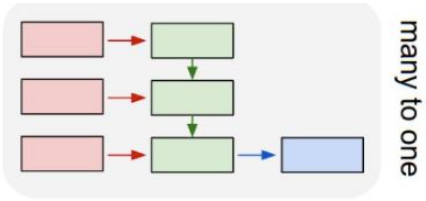
This movie is very slow in the beginning but picks up pace later on and has some great action sequences and comedy scenes.

Modeling Sequences



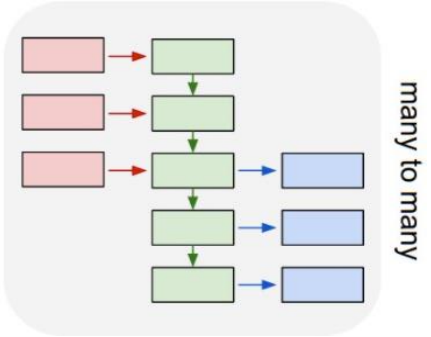
A person riding a motorbike on dirt road

Image Captioning

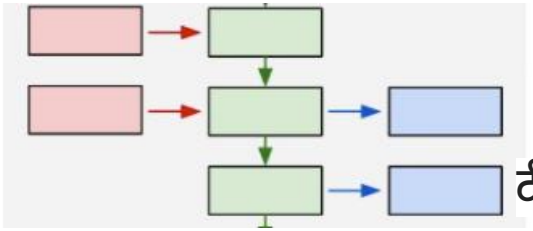


Awesome tutorial. Positive

Sentiment Analysis



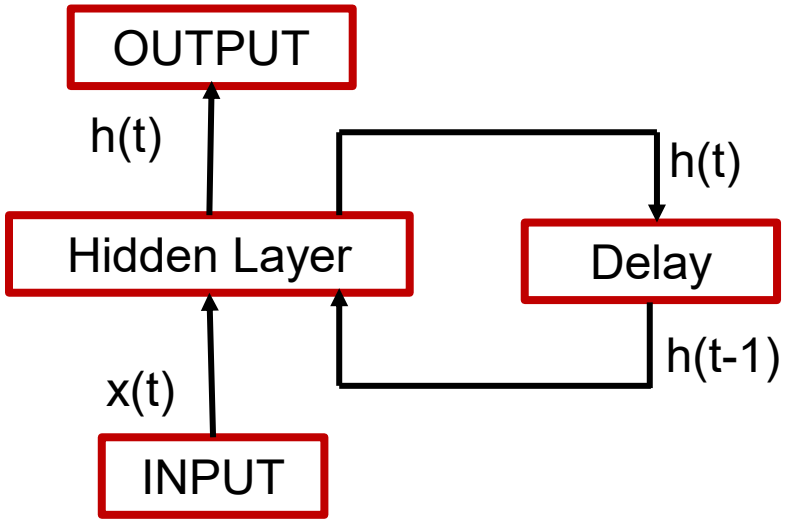
Happy Diwali



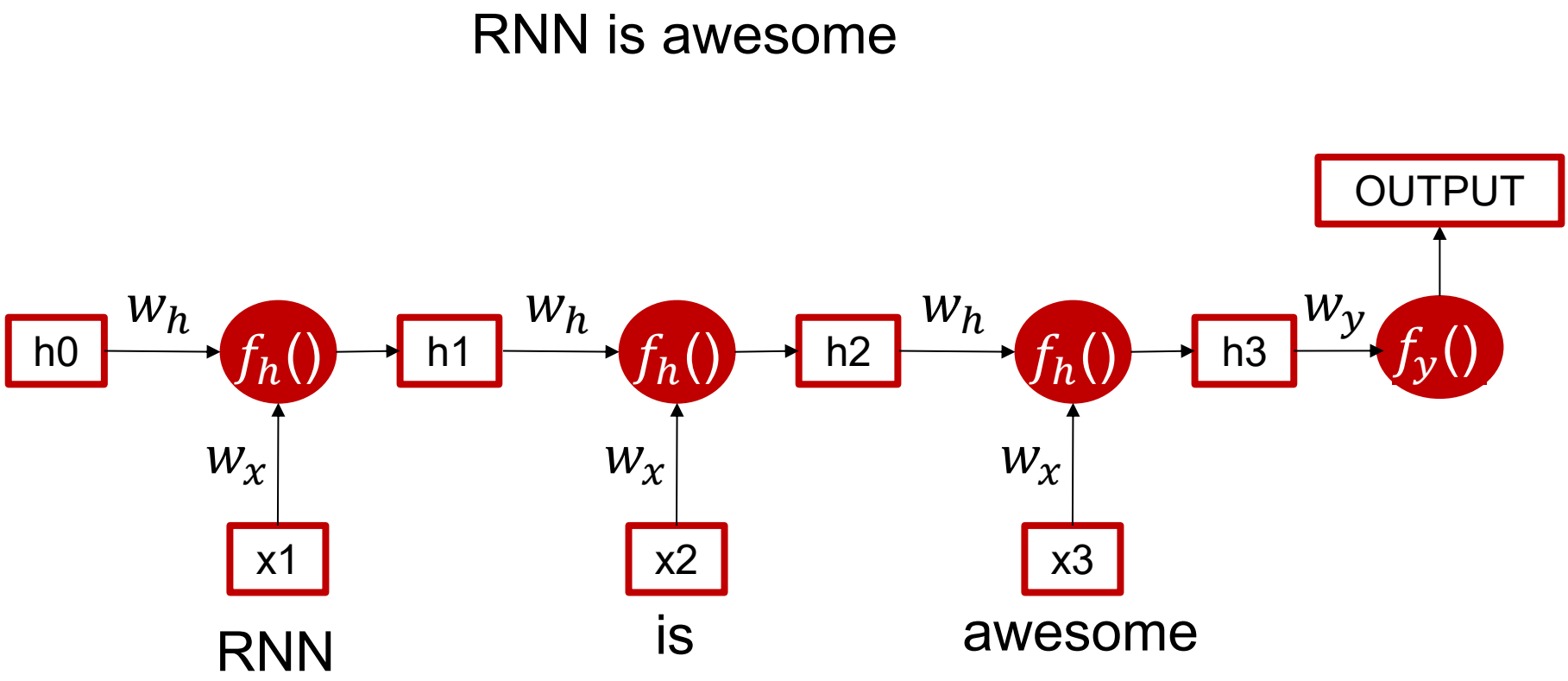
Machine Translation

What is RNN?

- Recurrent neural networks are connectionist models with the ability to selectively pass information across sequence steps, while processing sequential data one element at a time.
- Allows a memory of the previous inputs to persist in the model's internal state and influence the outcome.



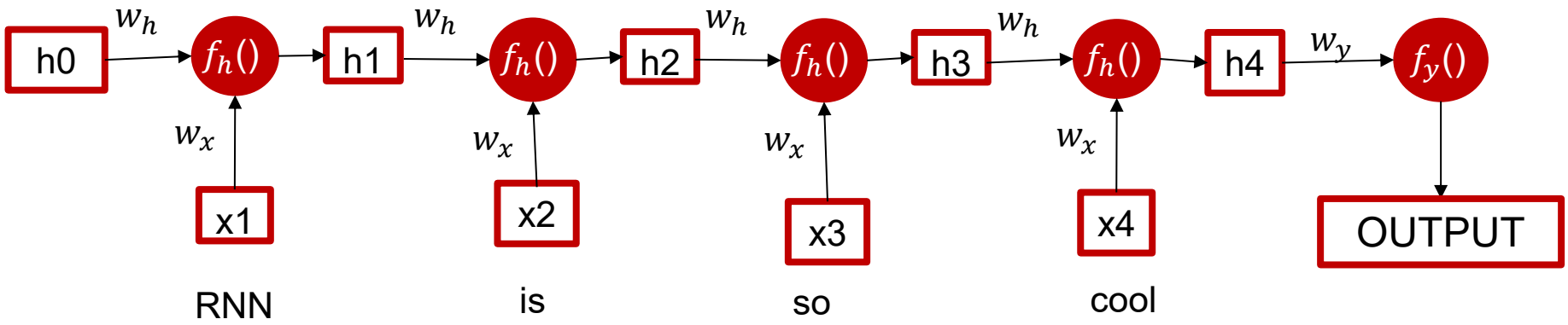
RNN (rolled over time)



$$h(t) = f_h\big(w_h * h(t-1) + w_x * x(t)\big)$$

RNN (rolled over time)

RNN is so cool



The Vanishing Gradient Problem

- RNN's use back propagation.
- Back propagation uses chain rule.
  - Chain rule multiplies derivatives
- If these derivatives are between 0 and 1 the product vanishes as the chain gets longer.
  - or the product explodes if the derivatives are greater than 1.
- Sigmoid activation function in RNN leads to this problem.
- Relu, in theory, avoids this problem but not in practice.

Problem with Vanishing or Exploding Gradients

- Don't allow us to learn long term dependencies.
  - Param is a hard worker.

VS.

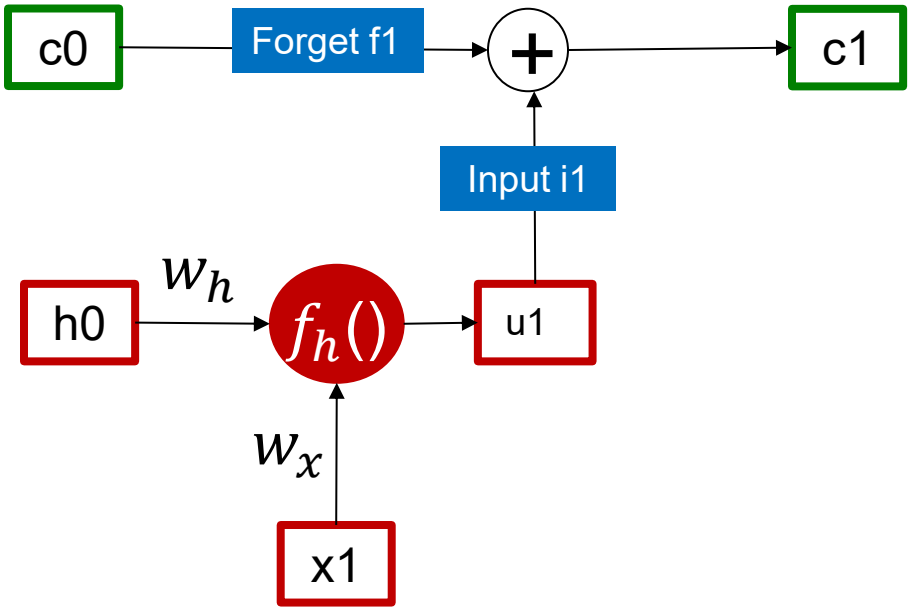
  - Param, student of Yong, is a hard worker.

BAD!!!!
Misguided!!!!
Unacceptable!!!!

Long Short Term Memory (LSTM)

- LSTM provide solution to the vanishing/exploding gradient problem.
- Solution: **Memory Cell**, which is updated at each step in the sequence.
- Three Gates control the flow of information to and from the Memory cell
  - Input Gate: protect the current step from irrelevant inputs
  - Output Gate: prevents current step from passing irrelevant information to later steps.
  - Forget Gate: limits information passed from one cell to the next.

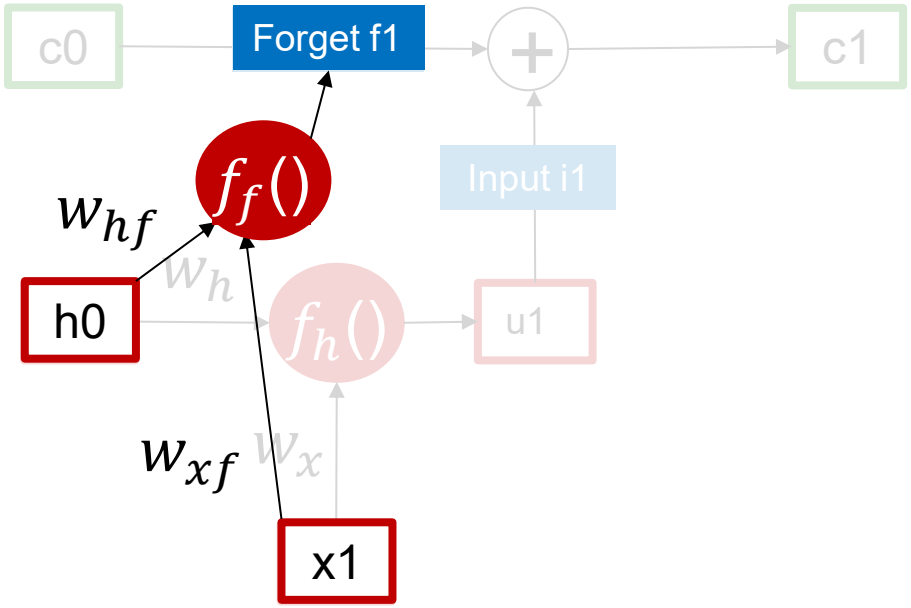
LSTM



$$c_1 = f_1 . c_0 + i_i . u_1$$

$$c_t = f_t . c_{t-1} + i_t . u_t$$

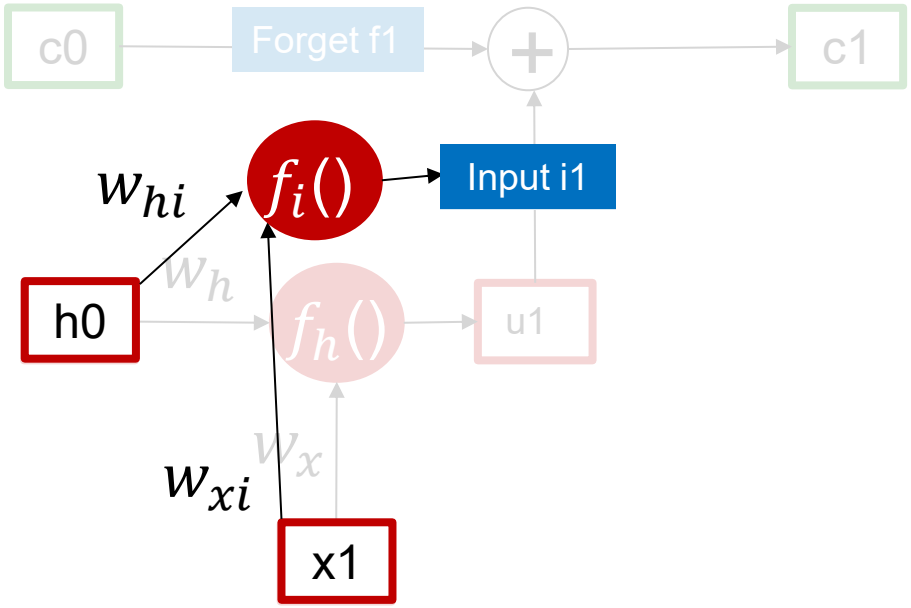
LSTM



$$f_1 = f_f\left(W_{hf} * h_0 + W_{xf} * x_1\right)$$

$$f_t = f_f\left(W_{hf} * h_{t-1} + W_{xf} * x_t\right)$$

LSTM



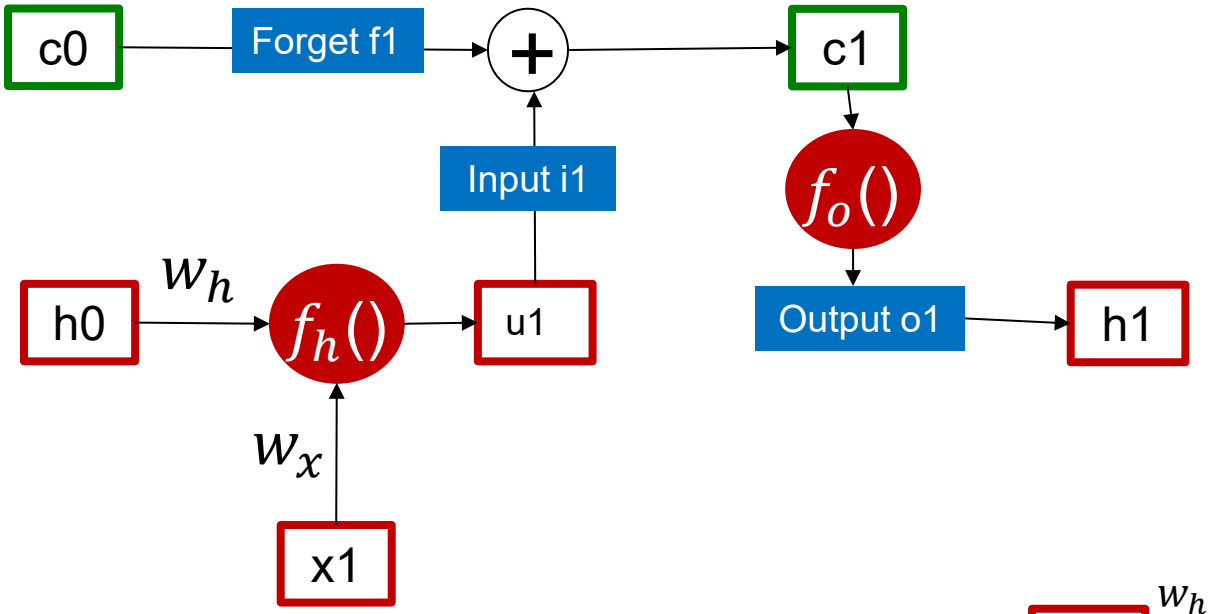
$$i_1 = f_i(W_{hi} * h_0 + W_{xi} * x_1)$$

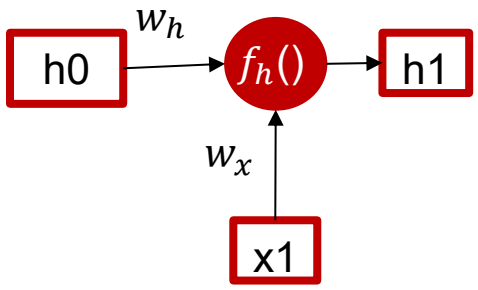
$$i_t = f_f(W_{hi} * h_{t-1} + W_{xi} * x_t)$$

LSTM

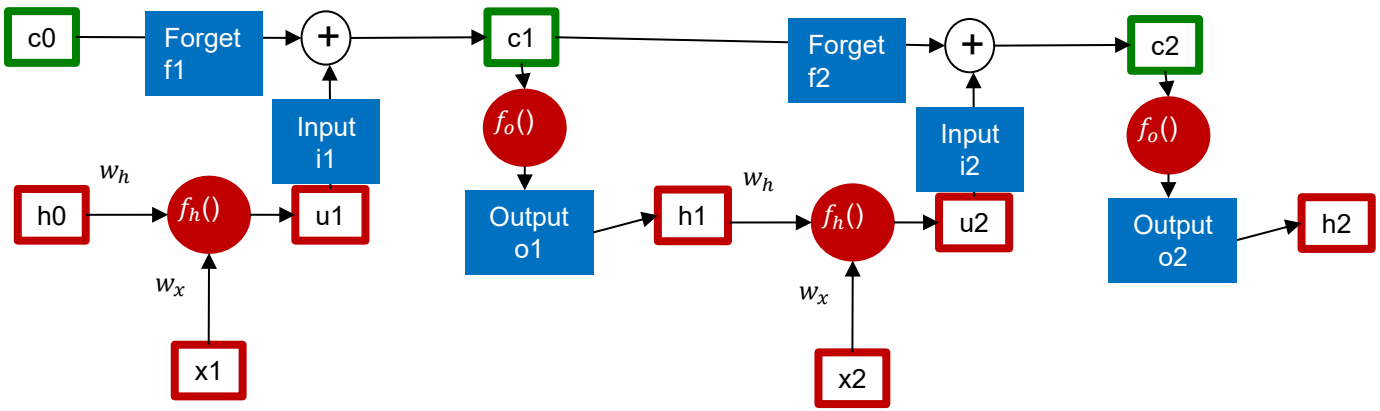
$$o_t = f_o(W_{ho} * h_{t-1} + W_{xo} * x_t)$$

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



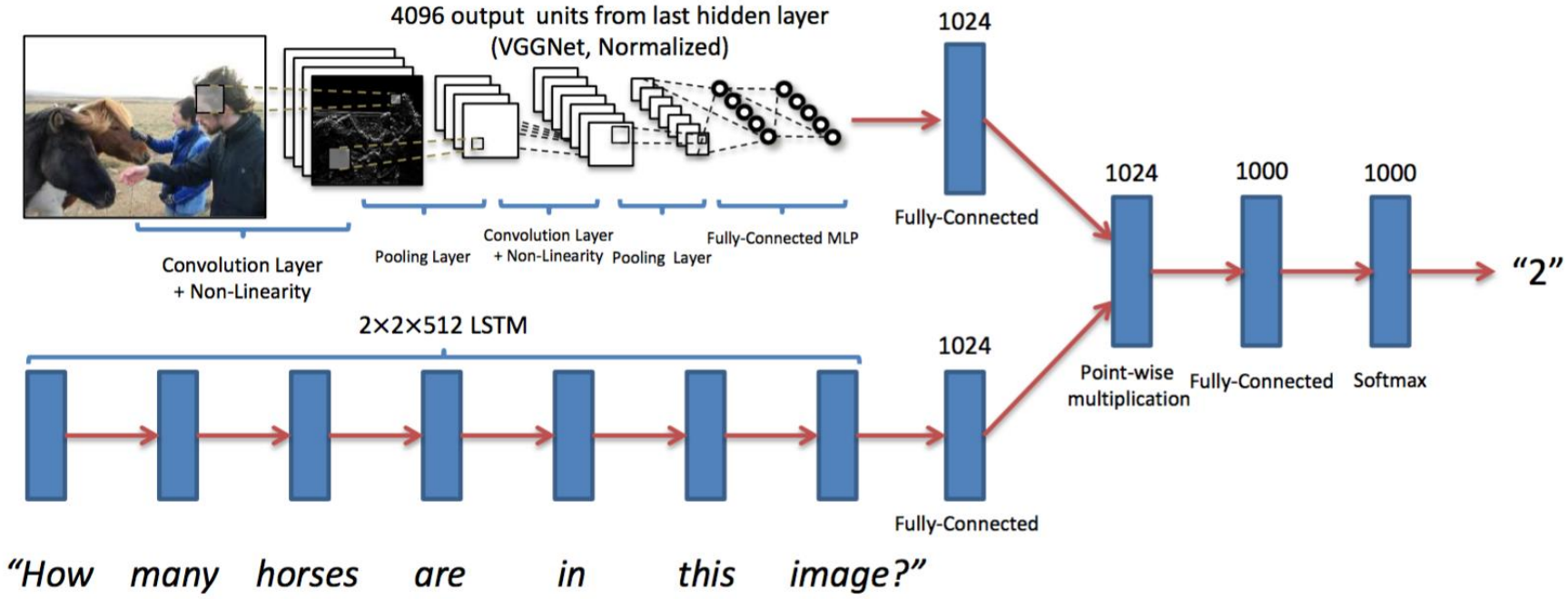


LSTM


c0
Forget
f1
+
Input
i1
c1
$f_o()$
Output
o1
Forget
f2
+
Input
i2
c2
$f_o()$
Output
o2
h0
w_h
$f_h()$
u1
w_x
x1
h1
w_h
$f_h()$
u2
w_x
x2
h2

Combining CNN and LSTM

Visual Question Answering



Thank you…

Any questions??

يرجى مسح رمز الاستجابة السريعة QR Code لتعبئة نموذج التغذية الراجعة حول المحاضرة. ملاحظاتكم مهمة لتحسين المحاضرات القادمة.

My google site