DSCI 572: Supervised Learning II

Lecture 1

Muhammad Abdul-Mageed

muhammad.mageed@ubc.ca

Deep Learning & NLP Lab

The University of British Columbia

Agenda

1. Overview of AI
2. Deep Learning
   - Biological Inspiration
   - Recurrent Neural Networks
   - Convolutional Neural Networks
   - ...

What is Artificial Intelligence?

Art . . .

Art *Contd.*

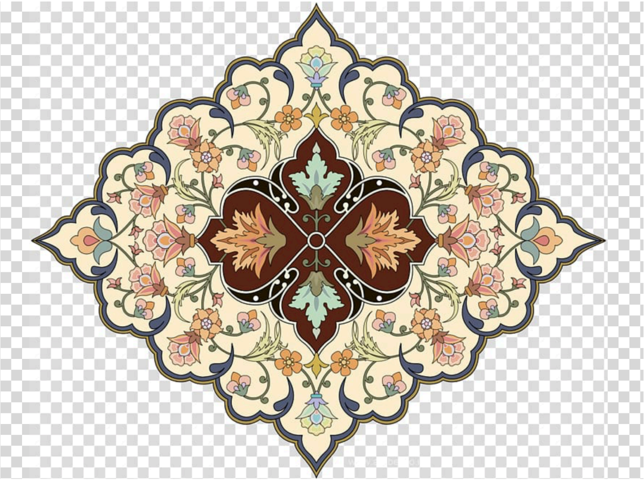


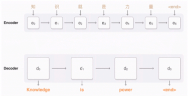
Figure: Legendary inventors in Greek mythology. Pygmalion and the Statue [Wikimedia commons].

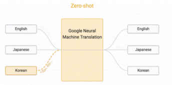
Can **programmable machines** become ***intelligent?***

Figure: **Ada Lovelace. 1842** [Wikipedia]

Deep Learning & Its Impact

Machine Translation

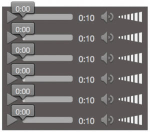


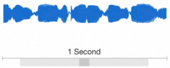


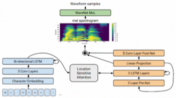
https://research.googleblog.com/

https://translator.microsoft.com/help/articles/neural/

Speech

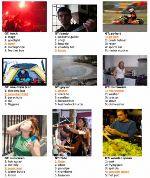






https://research.googleblog.com/

Image (Captioning)



[Microsoft Research: He et al., 2015]

https://research.googleblog.com/

Overview

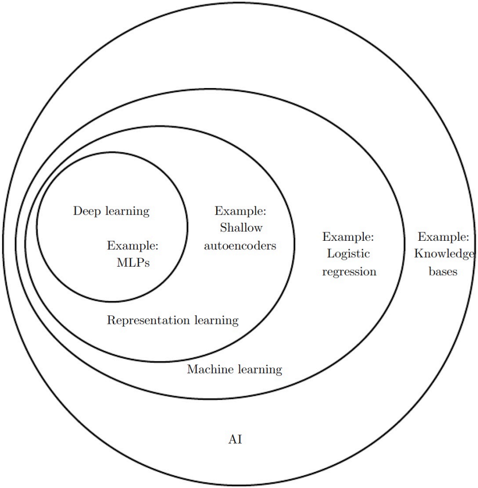


Figure: A Venn diagram situating DL within AI [Goodfellow et al., 2016].

Biological Inspiration

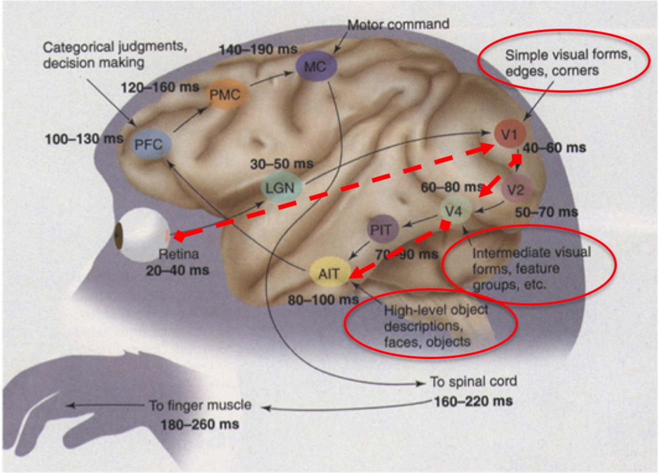


Figure: Visual cortex. [Modified from Simon Thorpe].

Networks with Multiple Layers

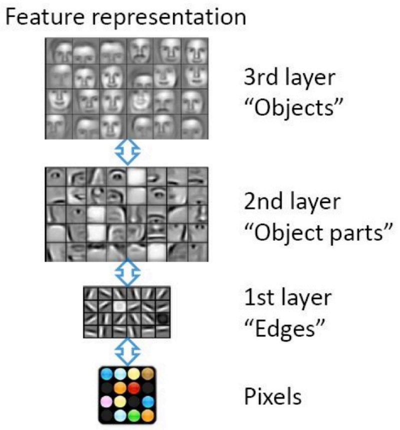


Figure: Leaerning in stages

Information processing in a neuron

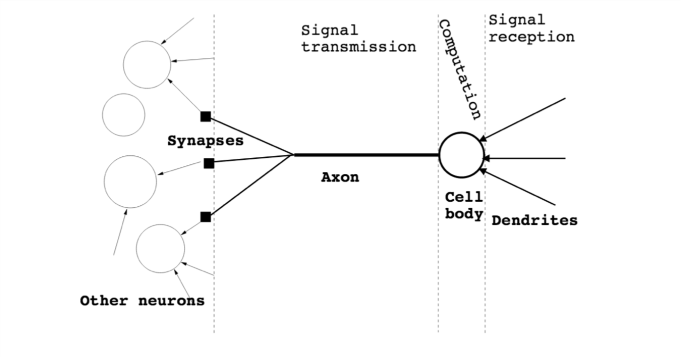


Figure: Information processing in a neuron. [Hyvarinen, Hurri, & Hoyer, 2009]

Neuron With Input

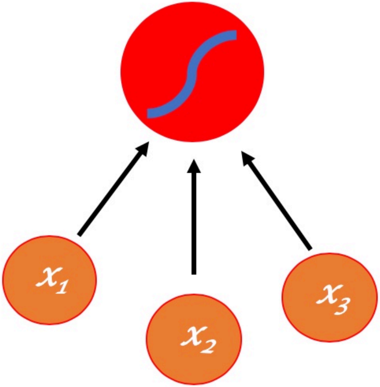


Figure: Input is a vector of **x** with three units, each taking an index j

A Vector For Connection Weights

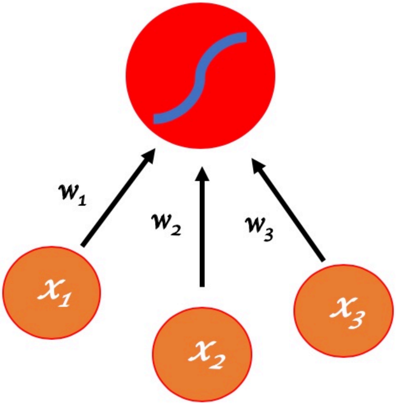


Figure: A vector of free parameters **w** with several items, each taking an index i

A Bias Unit

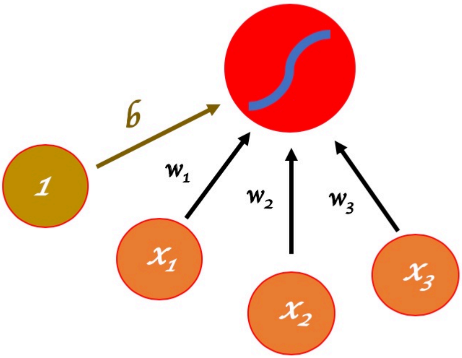


Figure: A bias unit b, equal to 1

A Deep Neural Network

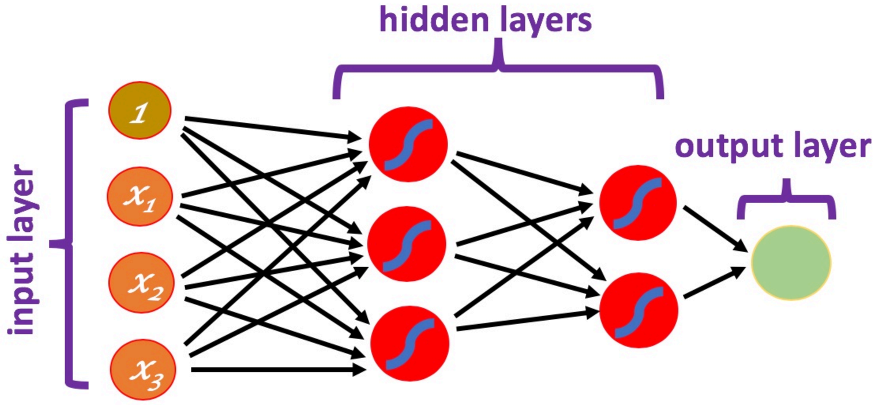


Figure: A **deep neural network**. The network has **2 hidden layers**.

RNN With Output

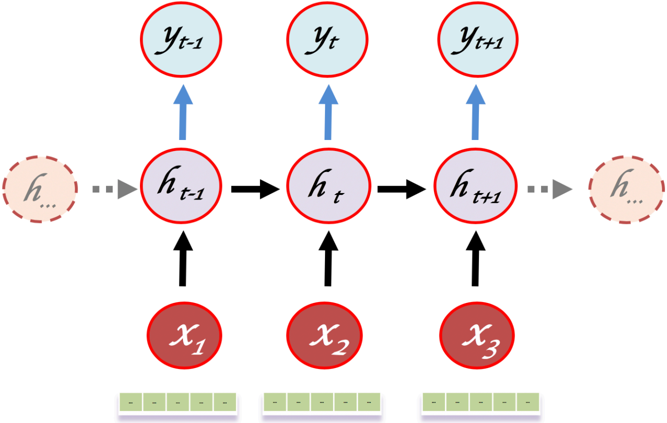


Figure: RNN is a generative model (can output a y at each time step).

Bidirectional RNNs

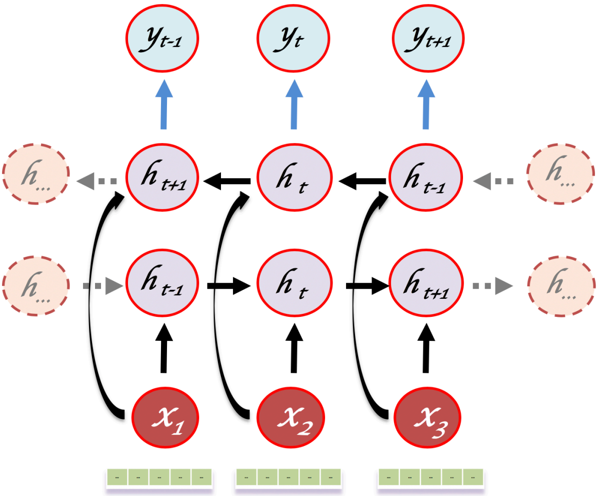


Figure: A bidirectional RNN (BiRNN).

Seq2Seq Models

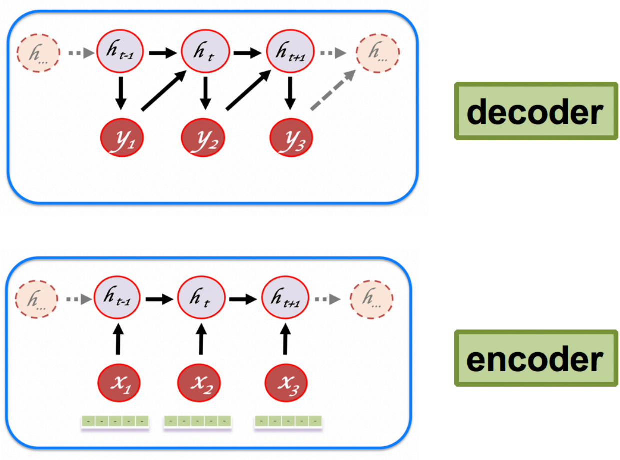


Figure: An Encoder and decoder.

Seq2Seq Models has Encoder Context

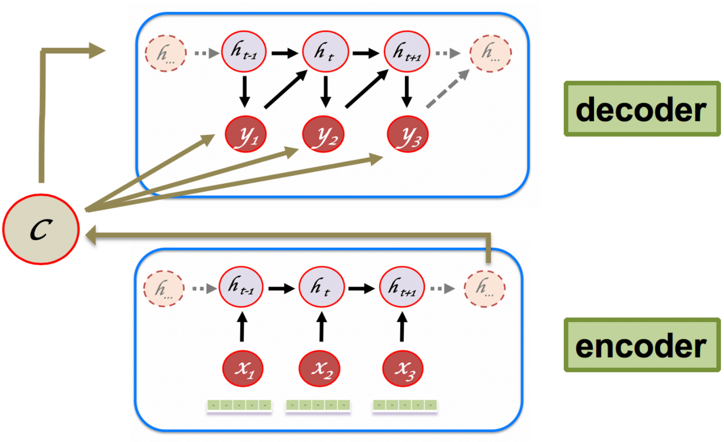
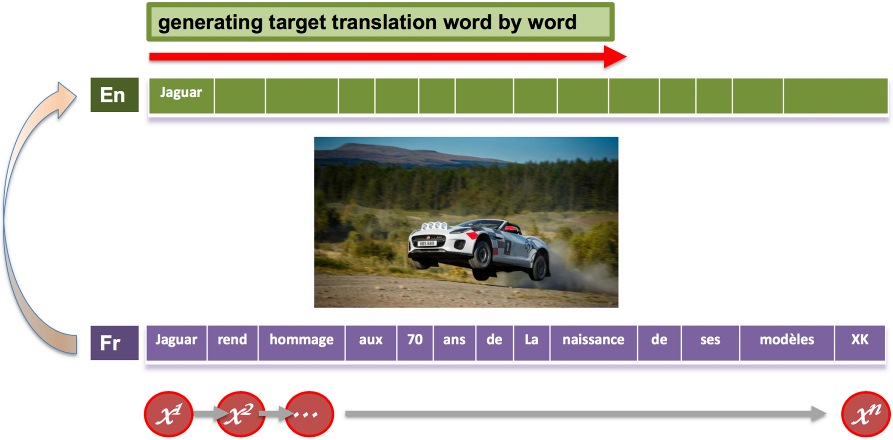
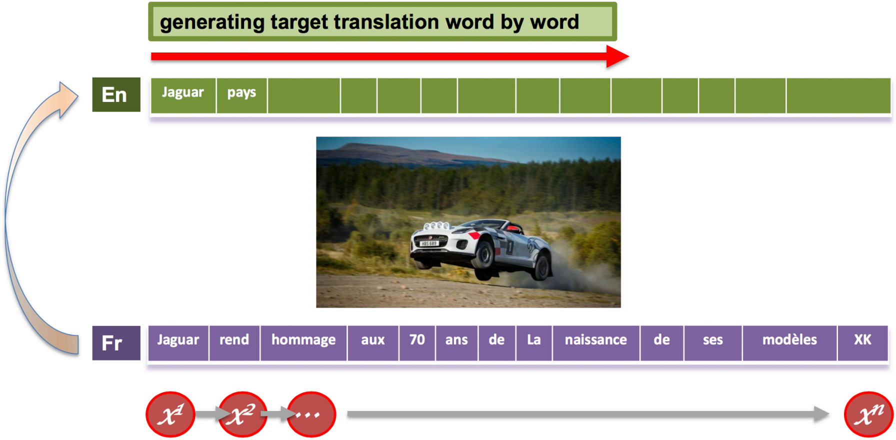


Figure: An Encoder and decoder connected via a context vector C.

Generating Target (En) One Word at a Time



Generating Target (En) One Word at a Time *Contd. I*



Generating Target (En) One Word at a Time *Contd. II*

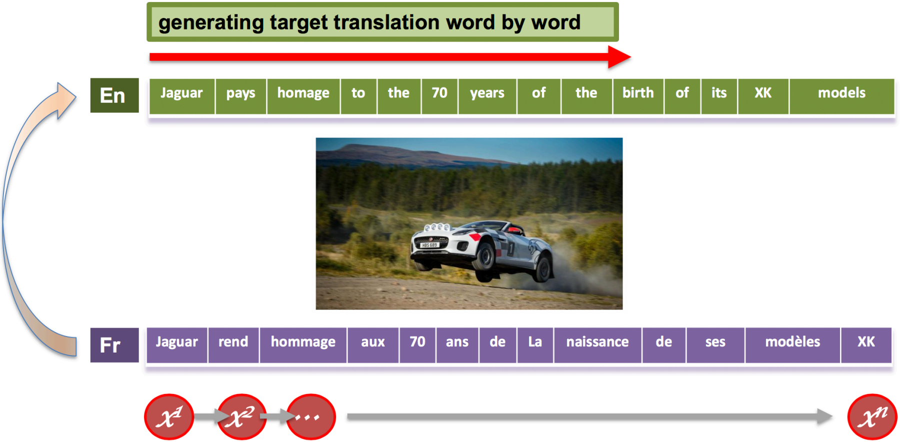
generating target translation word by word

En	Jaguar	pays	homage										

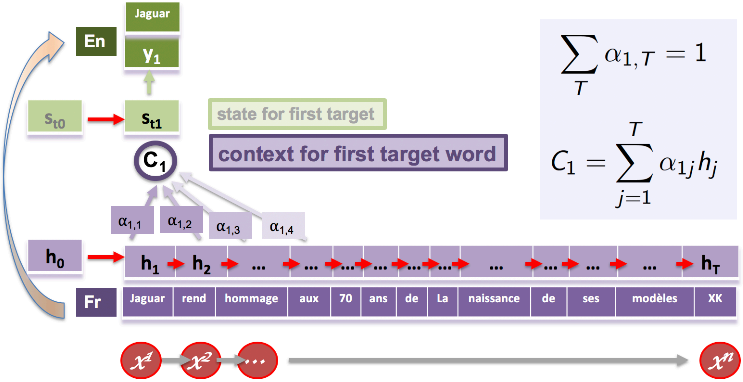
Fr	Jaguar	rend	hommage	aux	70	ans	de	La	naissance	de	ses	modèles	XK

$x^1 \rightarrow x^2 \rightarrow \dots \rightarrow x^n$

Generating Target (En) One Word at a Time *Contd. III*



Context For State One



$$\sum_T \alpha_{1,T} = 1$$

$$C_1 = \sum_{j=1}^{T} \alpha_{1j} h_j$$

Context Ingredients

- α: **attention weights**. (*We will learn these weights*)
- h: **activation from source**

5 x 5 matrix (e.g., input image)

1	1	0	0	0
0	0	1	1	1
0	0	0	0	0
1	1	0	0	1
1	1	0	1	1

3 x 3 kernel (*aka* filter)

1	0	0
0	1	1
1	0	0

Desired feature map

Figure: **Input**: 5 x 5 matrix/2D Tensor. We will convolve with a 3 x 3 kernel, with a `stride` of 1. Our goal is to acquire an feature map (filling in the 3 x 3 matrix on right-hand side).

Convolving 1

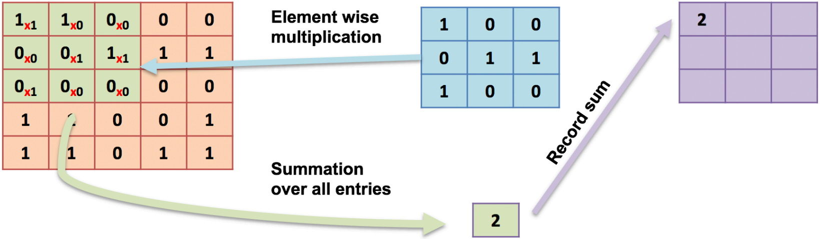


Figure: Filling in the first cell.

Convolving 2

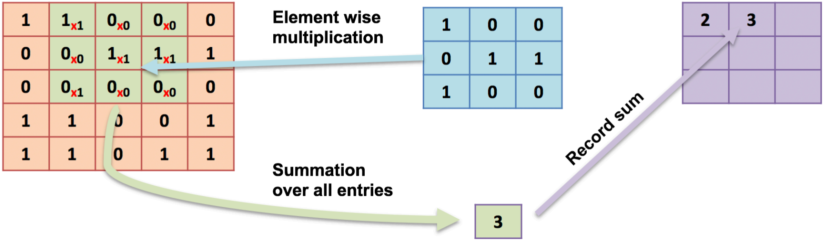


Figure: Filling in the second cell.

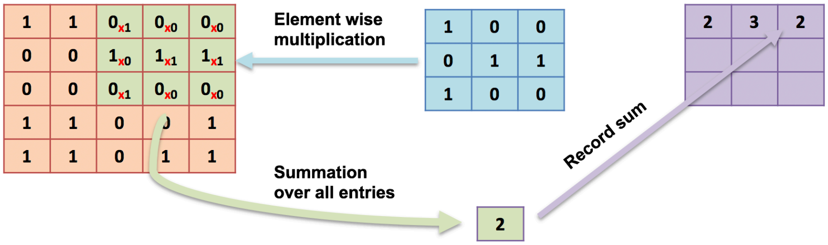


Figure: Filling in the third cell.

Convolving 4

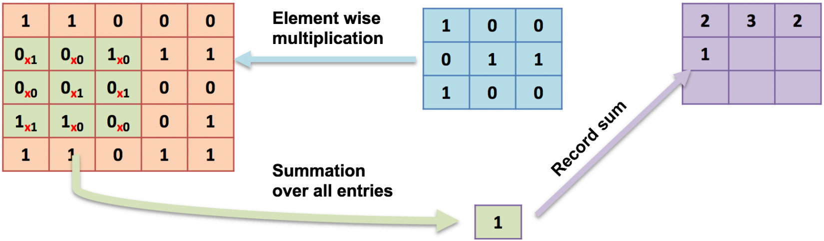


Figure: Filling in the fourth cell.

Convolving 5-9

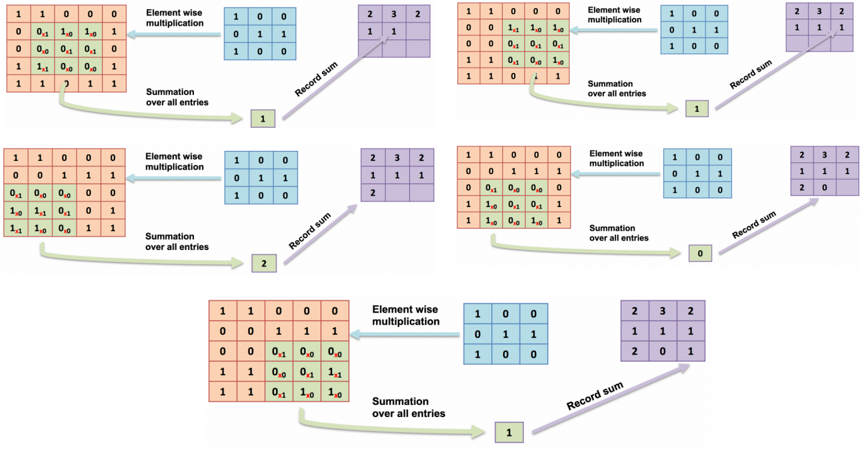


Figure: Filling in the fifth-to-ninth cells.

Convolving With 10 Kernels

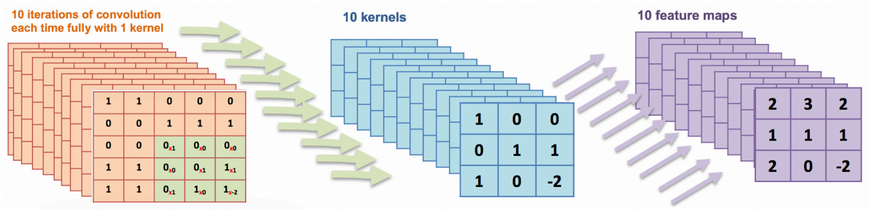


Figure: Convolving with 10 kernels, each with different values, we acquire 10 different feature maps.

Non-Linearity

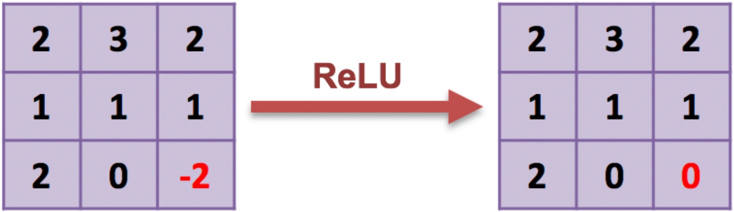


Figure: We actually apply a non-linear function like ReLU on the output of convolution operations such that we end up with meaningful feature maps.

Convolution With Non-Linearity

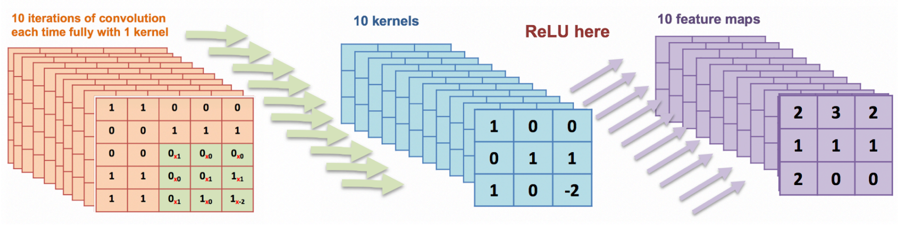


Figure: We actually apply a non-linear function like ReLU on the output of convolution operations such that we end up with meaningful feature maps.

Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jakob Uszkoreit*
Google Research
usz@google.com

Llion Jones*
Google Research
llion@google.com

Aidan N. Gomez* †
University of Toronto
aidan@cs.toronto.edu

Łukasz Kaiser*
Google Brain
lukaszkaiser@google.com

Illia Polosukhin* ‡
illia.polosukhin@gmail.com

Transformer *Contd.*

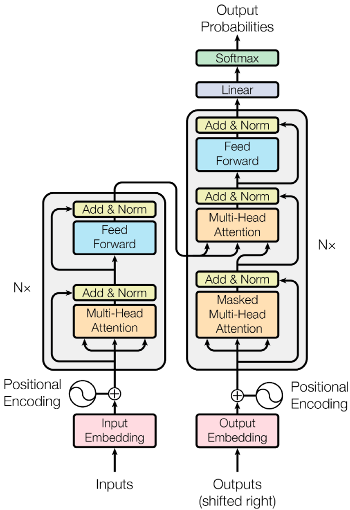


Figure: Transformer: Model Architecture.

(Variational) Autoencoders

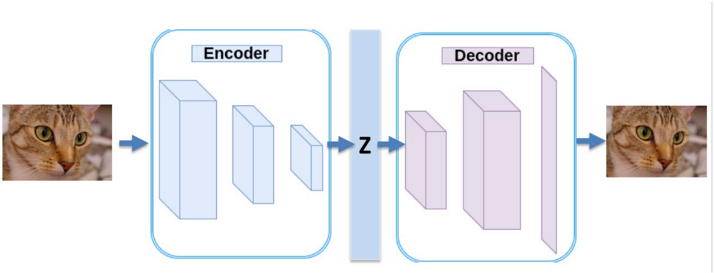


Figure: A variational autoencoder. It has an **encoder** (left) and a **decoder** (right).

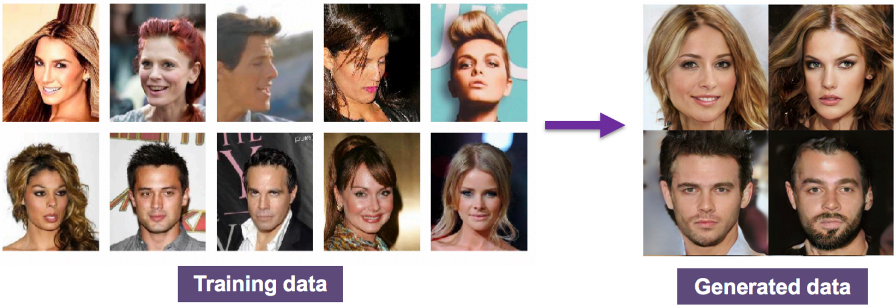


Figure: [From Karras et al, 2017; see paper]

GAN-Generated Images

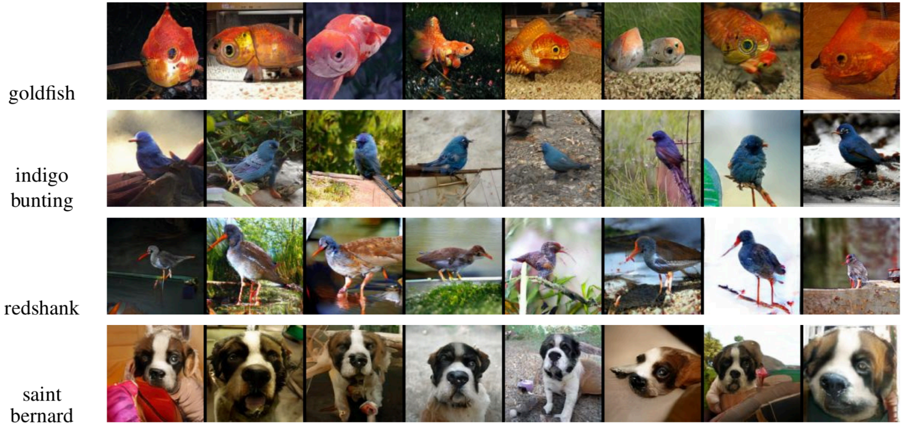


Figure: [From Zhang et al, 2018; see paper]

Thank You!

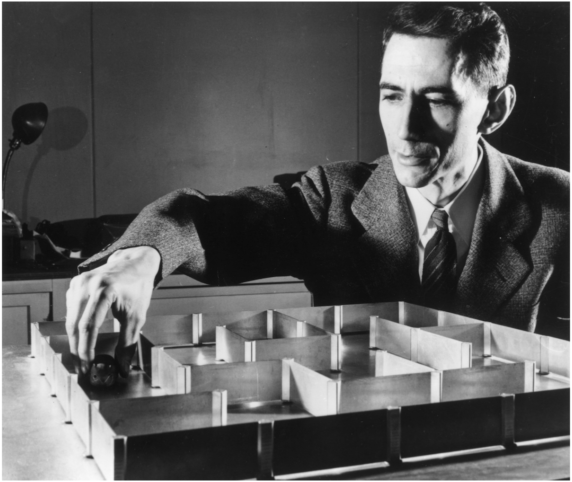


Figure: Claude Shannon. [From Time]