

# Deep Learning Basic

Lecture 1: Historical Review

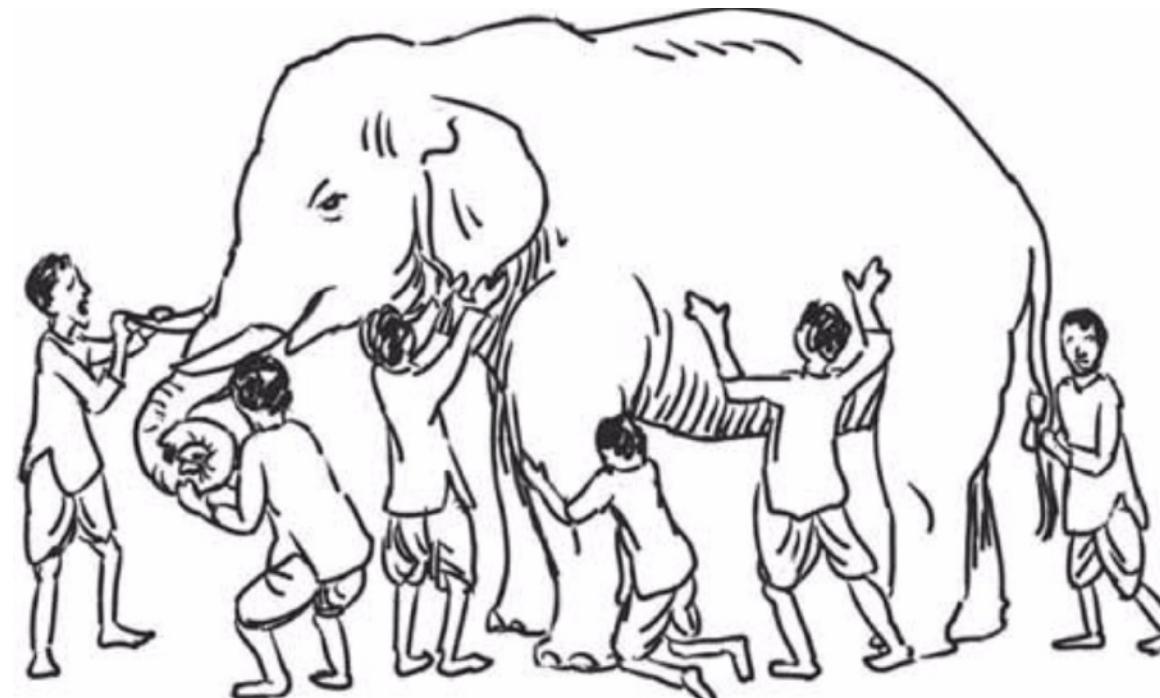
최성준 (고려대학교 인공지능학과)

# Introduction

# Introduction

---

- Disclaimer



<https://williepietersen.com/the-lessons-of-the-blind-men-and-the-elephant-2/>

# Introduction

- What make you a good deep learner?



# Implementation Skills

└ 요즘은 pytorch  
설명서

$\sin(x+y) = \sin x \cos y + \cos x \sin y$        $(\ln|x|)' = x^{-1}$        $\frac{d}{dx} \frac{1}{x^2+1} = -\frac{2x}{(x^2+1)^2}$        $\sin x = 0,5$        $\int_{-\pi/2}^{\pi/2} dx = [\ln|x| + \sqrt{x^2 + \omega^2}] \Big|_{-\pi/2}^{\pi/2} = C$        $(a+b)^2 = a^2 + 2ab$   
 $\frac{dy}{dx} = \frac{dy}{dx} \cdot \frac{dx}{dx} = \frac{dy}{dx}$        $\frac{1}{\sin A} = \frac{6}{\sqrt{13}B}$        $e^{ix} = 1 + 0i$        $\overline{A} \cdot (\overline{B} \cdot \overline{C}) = y = kx + m$   
  
 $f(n, m, a, b) = x \in (-\infty, -2)$        $(e^x)' = e^x$        $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$        $\mathcal{P} = \left[ \begin{array}{c|cc} & a_1 & a_2 \\ \hline a_1 & & \\ a_2 & & \end{array} \right] = \left[ \begin{array}{c|cc} & a_1 & a_2 \\ \hline a_1 & & \\ a_2 & & \end{array} \right] + \left[ \begin{array}{c|cc} & a_1 & a_2 \\ \hline a_1 & & \\ a_2 & & \end{array} \right]$   
 $a^2 + b^2 + c^2 = 2bc \cos A$   
  
 $y = \sin x$        $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 + b_1 & b_2 \\ a_2 & a_3 + b_1 & b_3 \\ a_3 & a_1 + b_1 & b_3 \end{pmatrix}$   
 $D = b^2 - 4ac$   
 $\frac{d}{dx^2} = 2$        $e^{2x} = e^x \cdot e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$   
  
 $\sin x = \operatorname{Im}[e^{ix}]$   
  
 $\cosh(x) = \frac{e^x + e^{-x}}{2}$   
 $X = 1$   
 $X' = 1 - 2 - X = \frac{1-2}{1-X}$   
 $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$   
 $a \cap b = \emptyset$   
 $\log_a \frac{\sin x}{\cos x} = \log_a \frac{\sqrt{1-x^2}}{\sqrt{1-x^2}} = 0$   
 $\cos(x+y) = \cos x \cos y - \sin x \sin y$   
  
 $\frac{1}{n} \sum_{i=1}^n f(x_i) \Delta x = \frac{1}{n} \sum_{i=1}^n f(\xi_i) \Delta x$   
 $\int_{-1}^1 f(x) dx = \int_{-1}^1 g(x) dx$   
 $\int_{-1}^1 x^2 dx = \frac{1}{3} x^3 \Big|_{-1}^1 = \frac{2}{3}$   
 $\cos 2x = 2 \cos x - 1$   
  
 $S = \frac{1}{2} ab \sin \alpha$   
 $y = x^2$   
  
 $\int_a^b f(x) dx = \int_a^b g(x) dx$   
 $\ln(a-b) = \int_a^b \frac{1}{x-a} dx$   
 $e^x = \sum_{n=0}^{\infty} \frac{f^n(x)}{n!} (x-a)^n$   
 $\cos x = \operatorname{Re}[e^{ix}]$   
 $x'_1 = 1$   
 $\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k$   
 $\frac{\partial}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}$   
 $\operatorname{tg} \alpha = \frac{\sin \alpha}{\cos \alpha}$   
 $(e^x)' = e^x$   
 $\sin x = -i \sin(ix)$   
 $\left(\frac{X'}{X}\right)' = \frac{1}{X^2}$   
 $\forall x > 0 \exists N \in \mathbb{N} \forall n > N |x_n| < \epsilon$   
 $(\ln x)' = \frac{1}{x}$   
  
 $\log_a(xy) = \log_a x + \log_a y$   
  
 $\sqrt{2} = 1.41$   
 $\sum$   
 $\pi = 3.14$   
 $y = x^2$   
 $\int_a^b f(x) dx$   
 $\int_a^b g(x) dx$   
 $\int_a^b h(x) dx$   
 $\int_a^b i(x) dx$   
 $\int_a^b j(x) dx$   
 $\int_a^b k(x) dx$   
 $\int_a^b l(x) dx$   
 $\int_a^b m(x) dx$   
 $\int_a^b n(x) dx$   
 $\int_a^b o(x) dx$   
 $\int_a^b p(x) dx$   
 $\int_a^b q(x) dx$   
 $\int_a^b r(x) dx$   
 $\int_a^b s(x) dx$   
 $\int_a^b t(x) dx$   
 $\int_a^b u(x) dx$   
 $\int_a^b v(x) dx$   
 $\int_a^b w(x) dx$   
 $\int_a^b x(x) dx$   
 $\int_a^b y(x) dx$   
 $\int_a^b z(x) dx$   
 $\int_a^b x^2(x) dx$   
 $\int_a^b y^2(x) dx$   
 $\int_a^b z^2(x) dx$   
 $\int_a^b x^3(x) dx$   
 $\int_a^b y^3(x) dx$   
 $\int_a^b z^3(x) dx$   
 $\int_a^b x^4(x) dx$   
 $\int_a^b y^4(x) dx$   
 $\int_a^b z^4(x) dx$   
 $\int_a^b x^5(x) dx$   
 $\int_a^b y^5(x) dx$   
 $\int_a^b z^5(x) dx$   
 $\int_a^b x^6(x) dx$   
 $\int_a^b y^6(x) dx$   
 $\int_a^b z^6(x) dx$   
 $\int_a^b x^7(x) dx$   
 $\int_a^b y^7(x) dx$   
 $\int_a^b z^7(x) dx$   
 $\int_a^b x^8(x) dx$   
 $\int_a^b y^8(x) dx$   
 $\int_a^b z^8(x) dx$   
 $\int_a^b x^9(x) dx$   
 $\int_a^b y^9(x) dx$   
 $\int_a^b z^9(x) dx$   
 $\int_a^b x^{10}(x) dx$   
 $\int_a^b y^{10}(x) dx$   
 $\int_a^b z^{10}(x) dx$   
 $\int_a^b x^{11}(x) dx$   
 $\int_a^b y^{11}(x) dx$   
 $\int_a^b z^{11}(x) dx$   
 $\int_a^b x^{12}(x) dx$   
 $\int_a^b y^{12}(x) dx$   
 $\int_a^b z^{12}(x) dx$   
 $\int_a^b x^{13}(x) dx$   
 $\int_a^b y^{13}(x) dx$   
 $\int_a^b z^{13}(x) dx$   
 $\int_a^b x^{14}(x) dx$   
 $\int_a^b y^{14}(x) dx$   
 $\int_a^b z^{14}(x) dx$   
 $\int_a^b x^{15}(x) dx$   
 $\int_a^b y^{15}(x) dx$   
 $\int_a^b z^{15}(x) dx$   
 $\int_a^b x^{16}(x) dx$   
 $\int_a^b y^{16}(x) dx$   
 $\int_a^b z^{16}(x) dx$   
 $\int_a^b x^{17}(x) dx$   
 $\int_a^b y^{17}(x) dx$   
 $\int_a^b z^{17}(x) dx$   
 $\int_a^b x^{18}(x) dx$   
 $\int_a^b y^{18}(x) dx$   
 $\int_a^b z^{18}(x) dx$   
 $\int_a^b x^{19}(x) dx$   
 $\int_a^b y^{19}(x) dx$   
 $\int_a^b z^{19}(x) dx$   
 $\int_a^b x^{20}(x) dx$   
 $\int_a^b y^{20}(x) dx$   
 $\int_a^b z^{20}(x) dx$   
 $\int_a^b x^{21}(x) dx$   
 $\int_a^b y^{21}(x) dx$   
 $\int_a^b z^{21}(x) dx$   
 $\int_a^b x^{22}(x) dx$   
 $\int_a^b y^{22}(x) dx$   
 $\int_a^b z^{22}(x) dx$   
 $\int_a^b x^{23}(x) dx$   
 $\int_a^b y^{23}(x) dx$   
 $\int_a^b z^{23}(x) dx$   
 $\int_a^b x^{24}(x) dx$   
 $\int_a^b y^{24}(x) dx$   
 $\int_a^b z^{24}(x) dx$   
 $\int_a^b x^{25}(x) dx$   
 $\int_a^b y^{25}(x) dx$   
 $\int_a^b z^{25}(x) dx$   
 $\int_a^b x^{26}(x) dx$   
 $\int_a^b y^{26}(x) dx$   
 $\int_a^b z^{26}(x) dx$   
 $\int_a^b x^{27}(x) dx$   
 $\int_a^b y^{27}(x) dx$   
 $\int_a^b z^{27}(x) dx$   
 $\int_a^b x^{28}(x) dx$   
 $\int_a^b y^{28}(x) dx$   
 $\int_a^b z^{28}(x) dx$   
 $\int_a^b x^{29}(x) dx$   
 $\int_a^b y^{29}(x) dx$   
 $\int_a^b z^{29}(x) dx$   
 $\int_a^b x^{30}(x) dx$   
 $\int_a^b y^{30}(x) dx$   
 $\int_a^b z^{30}(x) dx$   
 $\int_a^b x^{31}(x) dx$   
 $\int_a^b y^{31}(x) dx$   
 $\int_a^b z^{31}(x) dx$   
 $\int_a^b x^{32}(x) dx$   
 $\int_a^b y^{32}(x) dx$   
 $\int_a^b z^{32}(x) dx$   
 $\int_a^b x^{33}(x) dx$   
 $\int_a^b y^{33}(x) dx$   
 $\int_a^b z^{33}(x) dx$   
 $\int_a^b x^{34}(x) dx$   
 $\int_a^b y^{34}(x) dx$   
 $\int_a^b z^{34}(x) dx$   
 $\int_a^b x^{35}(x) dx$   
 $\int_a^b y^{35}(x) dx$   
 $\int_a^b z^{35}(x) dx$   
 $\int_a^b x^{36}(x) dx$   
 $\int_a^b y^{36}(x) dx$   
 $\int_a^b z^{36}(x) dx$   
 $\int_a^b x^{37}(x) dx$   
 $\int_a^b y^{37}(x) dx$   
 $\int_a^b z^{37}(x) dx$   
 $\int_a^b x^{38}(x) dx$   
 $\int_a^b y^{38}(x) dx$   
 $\int_a^b z^{38}(x) dx$   
 $\int_a^b x^{39}(x) dx$   
 $\int_a^b y^{39}(x) dx$   
 $\int_a^b z^{39}(x) dx$   
 $\int_a^b x^{40}(x) dx$   
 $\int_a^b y^{40}(x) dx$   
 $\int_a^b z^{40}(x) dx$   
 $\int_a^b x^{41}(x) dx$   
 $\int_a^b y^{41}(x) dx$   
 $\int_a^b z^{41}(x) dx$   
 $\int_a^b x^{42}(x) dx$   
 $\int_a^b y^{42}(x) dx$   
 $\int_a^b z^{42}(x) dx$   
 $\int_a^b x^{43}(x) dx$   
 $\int_a^b y^{43}(x) dx$   
 $\int_a^b z^{43}(x) dx$   
 $\int_a^b x^{44}(x) dx$   
 $\int_a^b y^{44}(x) dx$   
 $\int_a^b z^{44}(x) dx$   
 $\int_a^b x^{45}(x) dx$   
 $\int_a^b y^{45}(x) dx$   
 $\int_a^b z^{45}(x) dx$   
 $\int_a^b x^{46}(x) dx$   
 $\int_a^b y^{46}(x) dx$   
 $\int_a^b z^{46}(x) dx$   
 $\int_a^b x^{47}(x) dx$   
 $\int_a^b y^{47}(x) dx$   
 $\int_a^b z^{47}(x) dx$   
 $\int_a^b x^{48}(x) dx$   
 $\int_a^b y^{48}(x) dx$   
 $\int_a^b z^{48}(x) dx$   
 $\int_a^b x^{49}(x) dx$   
 $\int_a^b y^{49}(x) dx$   
 $\int_a^b z^{49}(x) dx$   
 $\int_a^b x^{50}(x) dx$   
 $\int_a^b y^{50}(x) dx$   
 $\int_a^b z^{50}(x) dx$   
 $\int_a^b x^{51}(x) dx$   
 $\int_a^b y^{51}(x) dx$   
 $\int_a^b z^{51}(x) dx$   
 $\int_a^b x^{52}(x) dx$   
 $\int_a^b y^{52}(x) dx$   
 $\int_a^b z^{52}(x) dx$   
 $\int_a^b x^{53}(x) dx$   
 $\int_a^b y^{53}(x) dx$   
 $\int_a^b z^{53}(x) dx$   
 $\int_a^b x^{54}(x) dx$   
 $\int_a^b y^{54}(x) dx$   
 $\int_a^b z^{54}(x) dx$   
 $\int_a^b x^{55}(x) dx$   
 $\int_a^b y^{55}(x) dx$   
 $\int_a^b z^{55}(x) dx$   
 $\int_a^b x^{56}(x) dx$   
 $\int_a^b y^{56}(x) dx$   
 $\int_a^b z^{56}(x) dx$   
 $\int_a^b x^{57}(x) dx$   
 $\int_a^b y^{57}(x) dx$   
 $\int_a^b z^{57}(x) dx$   
 $\int_a^b x^{58}(x) dx$   
 $\int_a^b y^{58}(x) dx$   
 $\int_a^b z^{58}(x) dx$   
 $\int_a^b x^{59}(x) dx$   
 $\int_a^b y^{59}(x) dx$   
 $\int_a^b z^{59}(x) dx$   
 $\int_a^b x^{60}(x) dx$   
 $\int_a^b y^{60}(x) dx$   
 $\int_a^b z^{60}(x) dx$   
 $\int_a^b x^{61}(x) dx$   
 $\int_a^b y^{61}(x) dx$   
 $\int_a^b z^{61}(x) dx$   
 $\int_a^b x^{62}(x) dx$   
 $\int_a^b y^{62}(x) dx$   
 $\int_a^b z^{62}(x) dx$   
 $\int_a^b x^{63}(x) dx$   
 $\int_a^b y^{63}(x) dx$   
 $\int_a^b z^{63}(x) dx$   
 $\int_a^b x^{64}(x) dx$   
 $\int_a^b y^{64}(x) dx$   
 $\int_a^b z^{64}(x) dx$   
 $\int_a^b x^{65}(x) dx$   
 $\int_a^b y^{65}(x) dx$   
 $\int_a^b z^{65}(x) dx$   
 $\int_a^b x^{66}(x) dx$   
 $\int_a^b y^{66}(x) dx$   
 $\int_a^b z^{66}(x) dx$   
 $\int_a^b x^{67}(x) dx$   
 $\int_a^b y^{67}(x) dx$   
 $\int_a^b z^{67}(x) dx$   
 $\int_a^b x^{68}(x) dx$   
 $\int_a^b y^{68}(x) dx$   
 $\int_a^b z^{68}(x) dx$   
 $\int_a^b x^{69}(x) dx$   
 $\int_a^b y^{69}(x) dx$   
 $\int_a^b z^{69}(x) dx$   
 $\int_a^b x^{70}(x) dx$   
 $\int_a^b y^{70}(x) dx$   
 $\int_a^b z^{70}(x) dx$   
 $\int_a^b x^{71}(x) dx$   
 $\int_a^b y^{71}(x) dx$   
 $\int_a^b z^{71}(x) dx$   
 $\int_a^b x^{72}(x) dx$   
 $\int_a^b y^{72}(x) dx$   
 $\int_a^b z^{72}(x) dx$   
 $\int_a^b x^{73}(x) dx$   
 $\int_a^b y^{73}(x) dx$   
 $\int_a^b z^{73}(x) dx$   
 $\int_a^b x^{74}(x) dx$   
 $\int_a^b y^{74}(x) dx$   
 $\int_a^b z^{74}(x) dx$   
 $\int_a^b x^{75}(x) dx$   
 $\int_a^b y^{75}(x) dx$   
 $\int_a^b z^{75}(x) dx$   
 $\int_a^b x^{76}(x) dx$   
 $\int_a^b y^{76}(x) dx$   
 $\int_a^b z^{76}(x) dx$   
 $\int_a^b x^{77}(x) dx$   
 $\int_a^b y^{77}(x) dx$   
 $\int_a^b z^{77}(x) dx$   
 $\int_a^b x^{78}(x) dx$   
 $\int_a^b y^{78}(x) dx$   
 $\int_a^b z^{78}(x) dx$   
 $\int_a^b x^{79}(x) dx$   
 $\int_a^b y^{79}(x) dx$   
 $\int_a^b z^{79}(x) dx$   
 $\int_a^b x^{80}(x) dx$   
 $\int_a^b y^{80}(x) dx$   
 $\int_a^b z^{80}(x) dx$   
 $\int_a^b x^{81}(x) dx$   
 $\int_a^b y^{81}(x) dx$   
 $\int_a^b z^{81}(x) dx$   
 $\int_a^b x^{82}(x) dx$   
 $\int_a^b y^{82}(x) dx$   
 $\int_a^b z^{82}(x) dx$   
 $\int_a^b x^{83}(x) dx$   
 $\int_a^b y^{83}(x) dx$   
 $\int_a^b z^{83}(x) dx$   
 $\int_a^b x^{84}(x) dx$   
 $\int_a^b y^{84}(x) dx$   
 $\int_a^b z^{84}(x) dx$   
 $\int_a^b x^{85}(x) dx$   
 $\int_a^b y^{85}(x) dx$   
 $\int_a^b z^{85}(x) dx$   
 $\int_a^b x^{86}(x) dx$   
 $\int_a^b y^{86}(x) dx$   
 $\int_a^b z^{86}(x) dx$   
 $\int_a^b x^{87}(x) dx$   
 $\int_a^b y^{87}(x) dx$   
 $\int_a^b z^{87}(x) dx$   
 $\int_a^b x^{88}(x) dx$   
 $\int_a^b y^{88}(x) dx$   
 $\int_a^b z^{88}(x) dx$   
 $\int_a^b x^{89}(x) dx$   
 $\int_a^b y^{89}(x) dx$   
 $\int_a^b z^{89}(x) dx$   
 $\int_a^b x^{90}(x) dx$   
 $\int_a^b y^{90}(x) dx$   
 $\int_a^b z^{90}(x) dx$   
 $\int_a^b x^{91}(x) dx$   
 $\int_a^b y^{91}(x) dx$   
 $\int_a^b z^{91}(x) dx$   
 $\int_a^b x^{92}(x) dx$   
 $\int_a^b y^{92}(x) dx$   
 $\int_a^b z^{92}(x) dx$   
 $\int_a^b x^{93}(x) dx$   
 $\int_a^b y^{93}(x) dx$   
 $\int_a^b z^{93}(x) dx$   
 $\int_a^b x^{94}(x) dx$   
 $\int_a^b y^{94}(x) dx$   
 $\int_a^b z^{94}(x) dx$   
 $\int_a^b x^{95}(x) dx$   
 $\int_a^b y^{95}(x) dx$   
 $\int_a^b z^{95}(x) dx$   
 $\int_a^b x^{96}(x) dx$   
 $\int_a^b y^{96}(x) dx$   
 $\int_a^b z^{96}(x) dx$   
 $\int_a^b x^{97}(x) dx$   
 $\int_a^b y^{97}(x) dx$   
 $\int_a^b z^{97}(x) dx$   
 $\int_a^b x^{98}(x) dx$   
 $\int_a^b y^{98}(x) dx$   
 $\int_a^b z^{98}(x) dx$   
 $\int_a^b x^{99}(x) dx$   
 $\int_a^b y^{99}(x) dx$   
 $\int_a^b z^{99}(x) dx$   
 $\int_a^b x^{100}(x) dx$   
 $\int_a^b y^{100}(x) dx$   
 $\int_a^b z^{100}(x) dx$

# Math Skills (Linear Algebra, Probability)

boostcamp aitech



# Knowing a lot of recent Papers

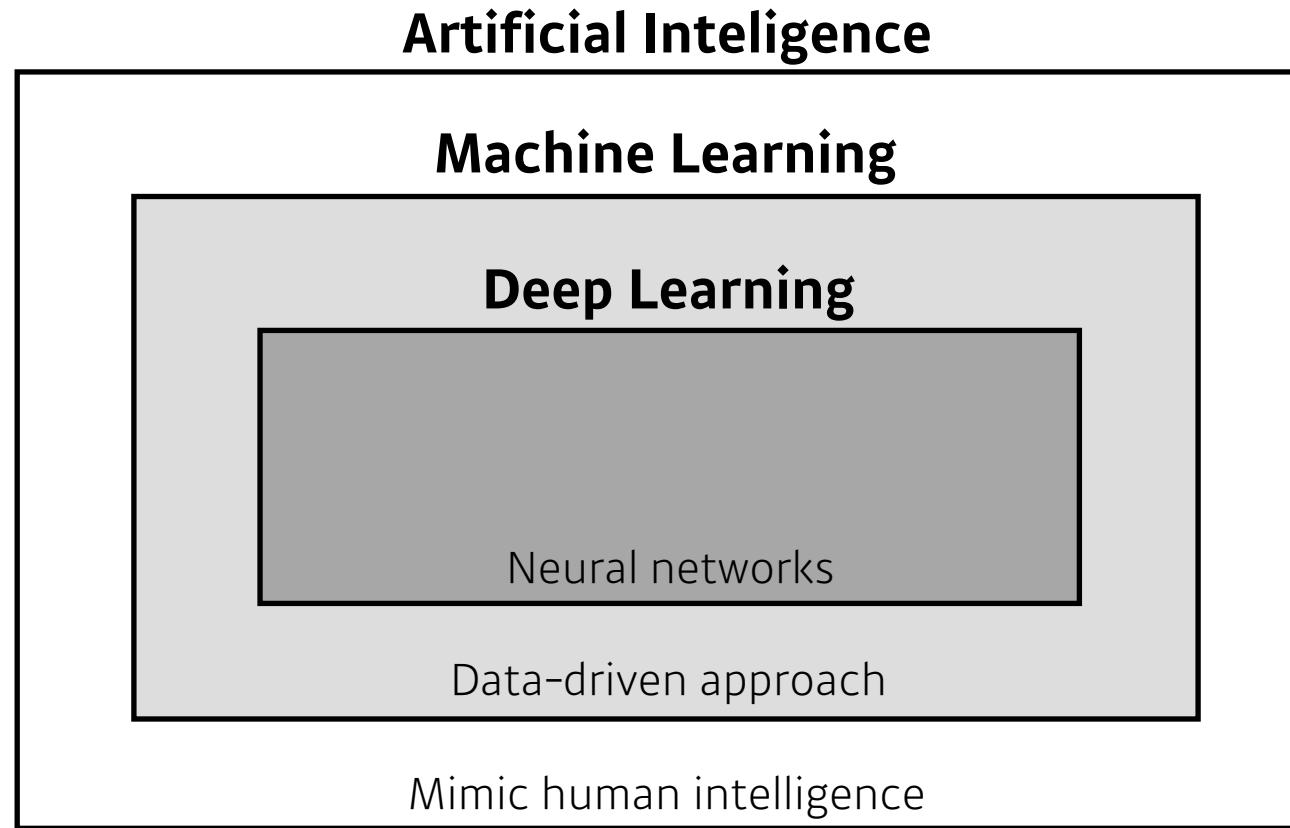
# Contents

---

1. Historical Review
2. Neural Networks & Multi-Layer Perceptron
3. Optimization Methods
4. Convolutional Neural Networks
5. Modern CNN
6. Computer Vision Applications
7. Recurrent Neural Networks
8. Transformer
9. Generative Models Part1
10. Generative Models Part2

# Introduction

---



# Introduction

---

- ➊ Key Components of Deep Learning
  - ➌ The data that the model can learn from
  - ➌ The model how to transform the data
  - ➌ The loss function that quantifies the badness of the model
  - ➌ The algorithm to adjust the parameters to minimize the loss

# Data

- Data depend on the type of the problem to solve.

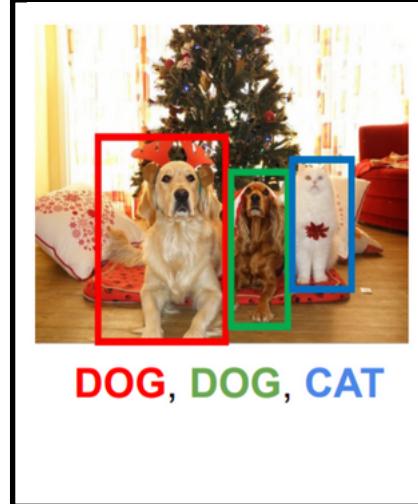
Classification



Semantic Segmentation



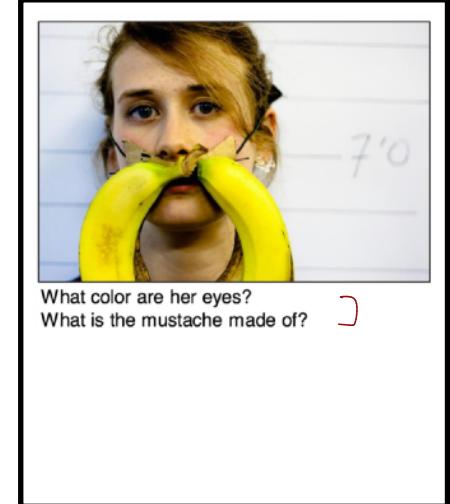
Detection



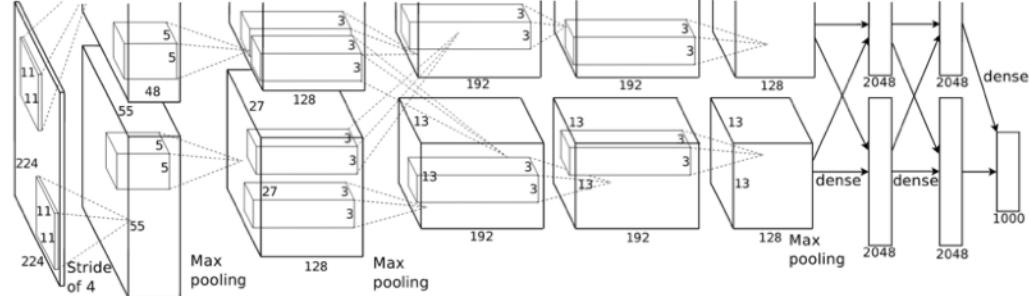
Pose Estimation



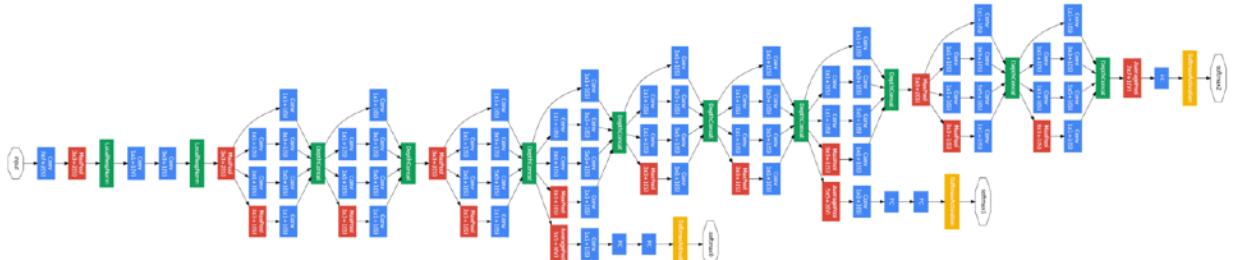
Visual QnA



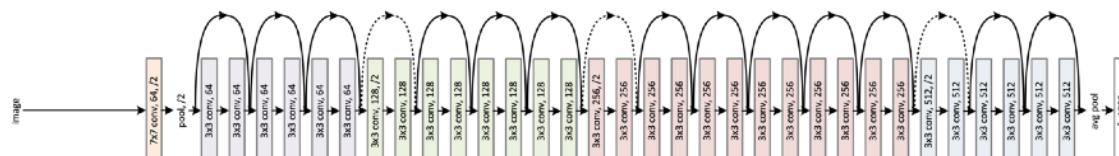
# Model



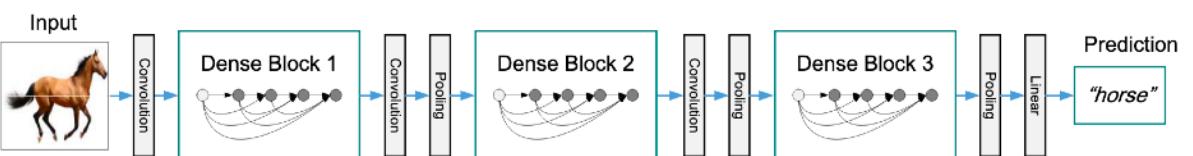
AlexNet



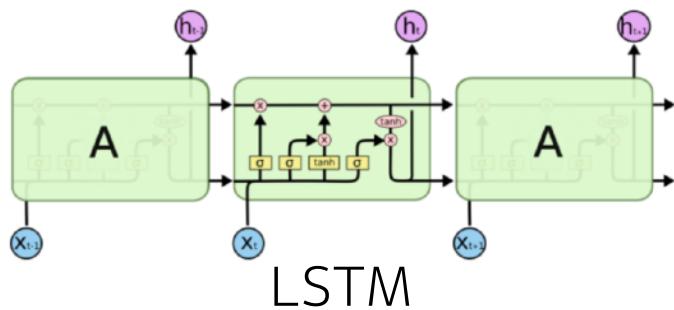
GoogLeNet



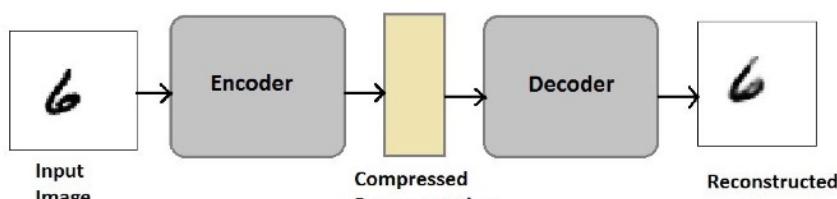
ResNet



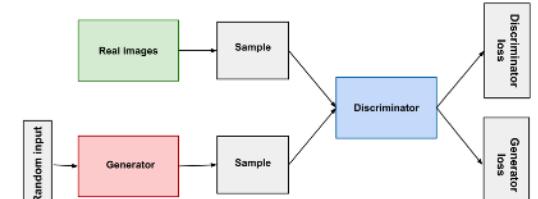
DenseNet



LSTM



Deep AutoEncoders



GAN

# LOSS

→ 이 낙제 점수는 문제의 궁극적인 목표입니다.

- The loss function is a proxy of what we want to achieve.

Regression Task

Classification Task

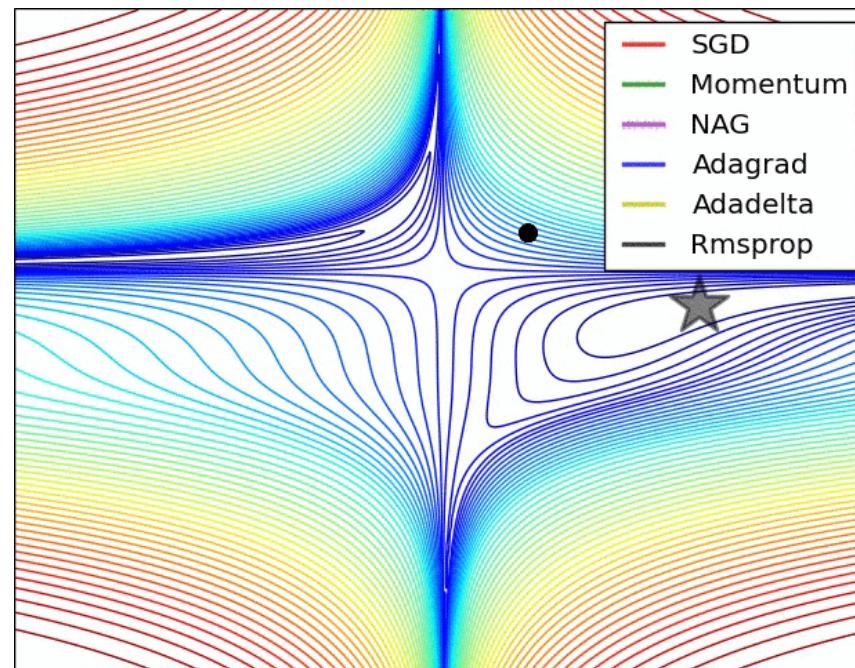
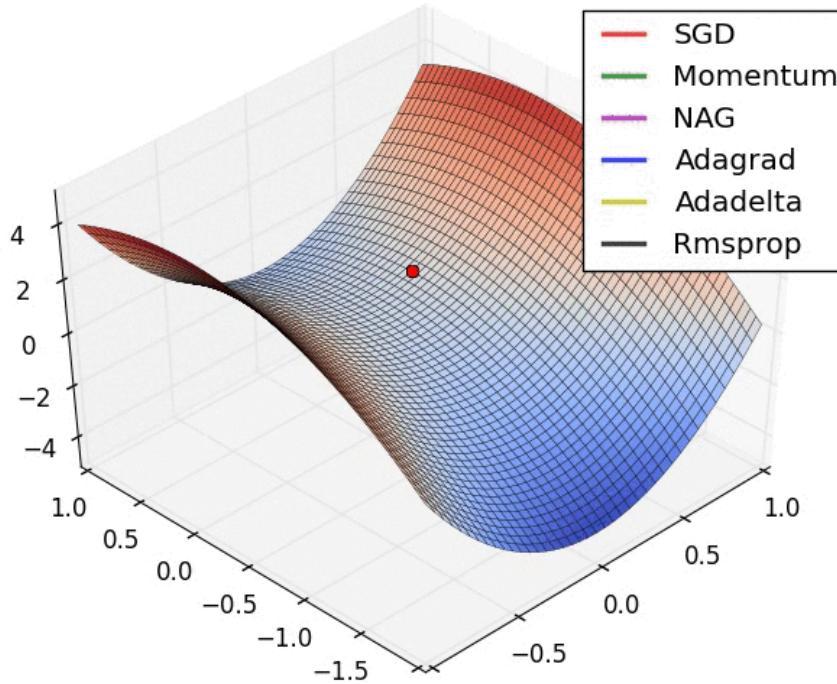
Probabilistic Task

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N \sum_{d=1}^D (y_i^{(d)} - \hat{y}_i^{(d)})^2$$

$$\text{CE} = -\frac{1}{N} \sum_{i=1}^N \sum_{d=1}^D y_i^{(d)} \log \hat{y}_i^{(d)}$$

$$\text{MLE} = \frac{1}{N} \sum_{i=1}^N \sum_{d=1}^D \log \mathcal{N}(y_i^{(d)}; \hat{y}_i^{(d)}, 1) \quad (= \text{MSE})$$

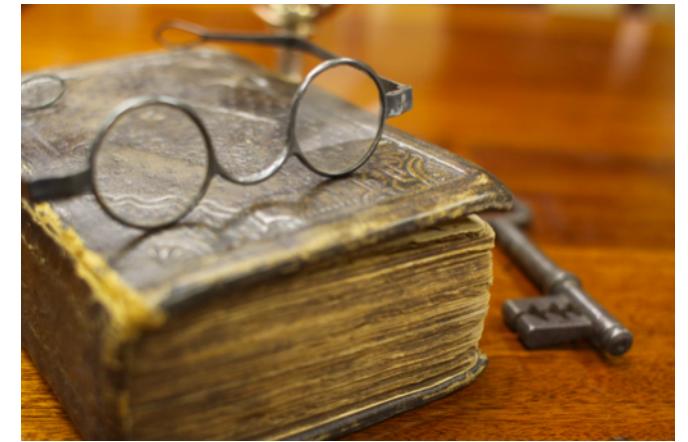
# Optimization Algorithm



과거/소 학습을 방지하기 위해

- Dropout
- Early stopping
- k-fold validation
- Weight decay
- Batch normalization
- MixUp
- Ensemble
- Bayesian Optimization

# Historical Review



<https://www.umc.org/en/who-we-are/history>

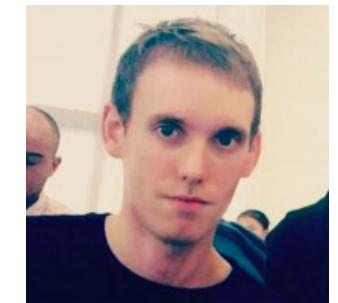
# Historical Review

---

## Deep Learning's Most Important Ideas - A Brief Historical Review

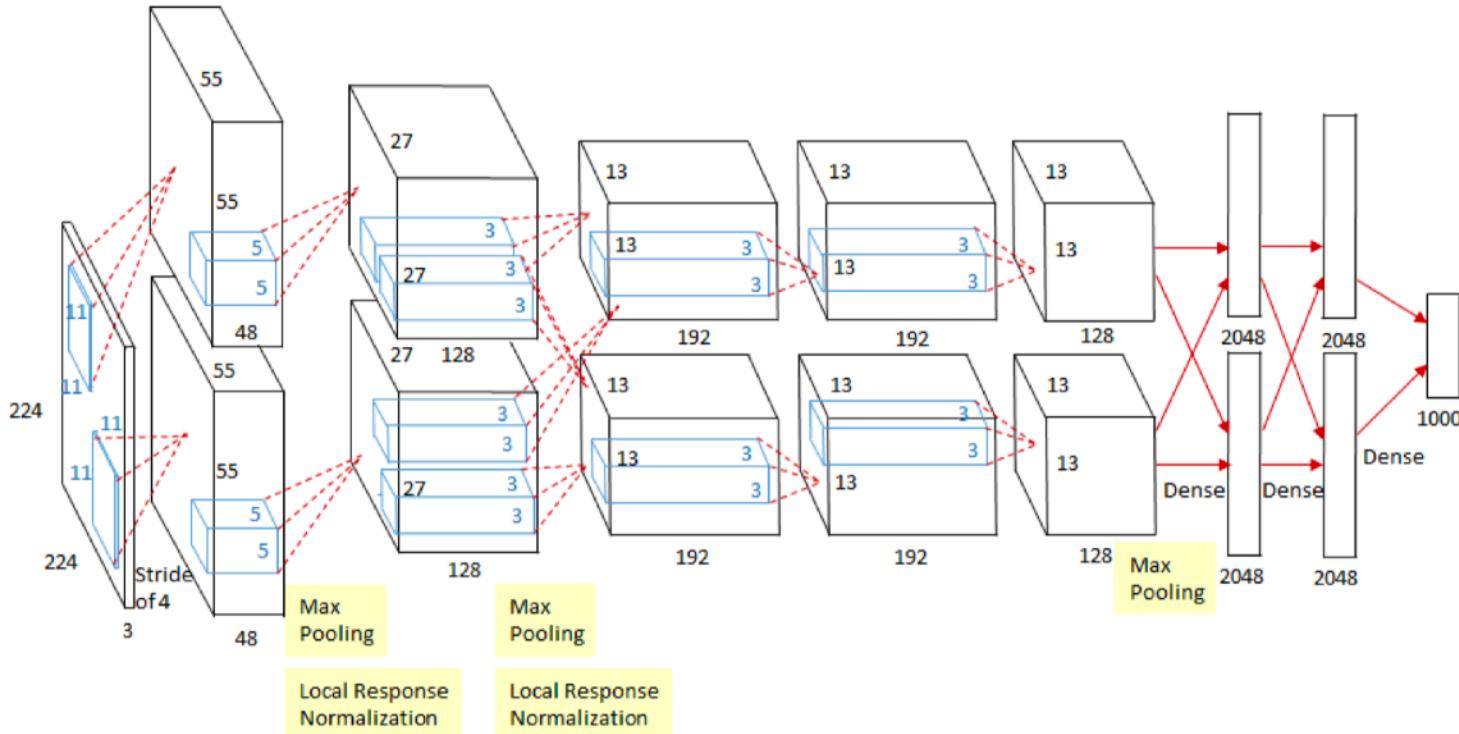
Denny Britz

2020-07-29



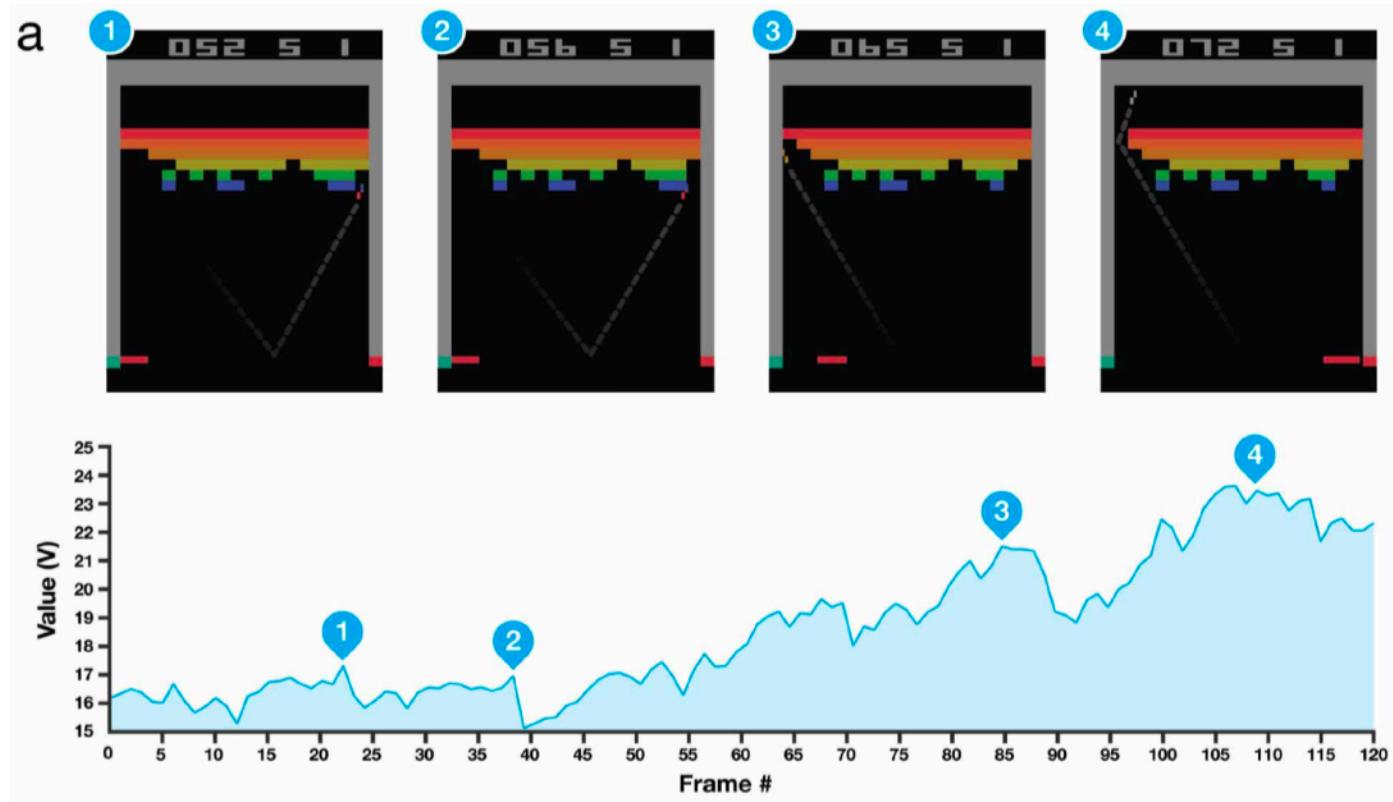
# 2012 - AlexNet

224 x 224 0.0171875 Gflops



# 2013 - DQN

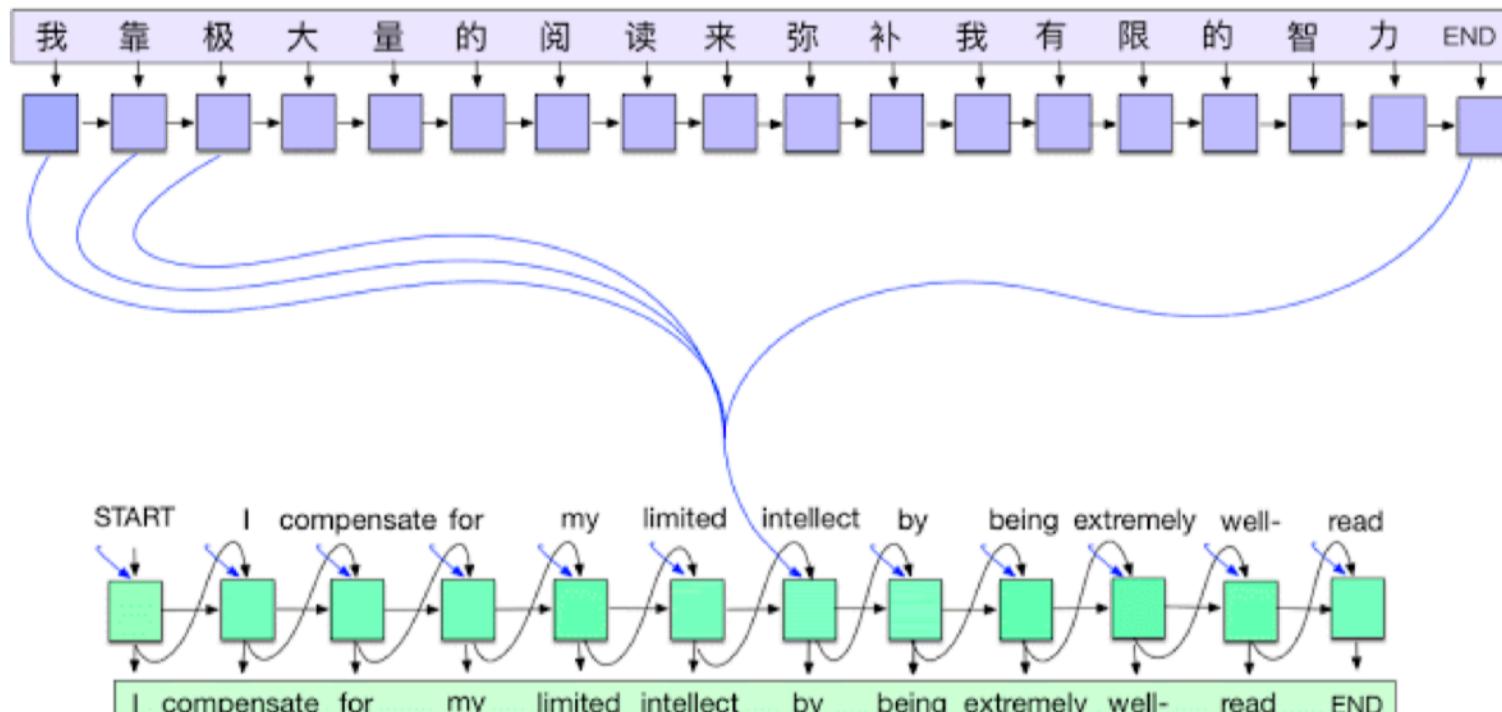
딥러닝을 이용한 학습 방법.



# 2014 - Encoder / Decoder

NMT

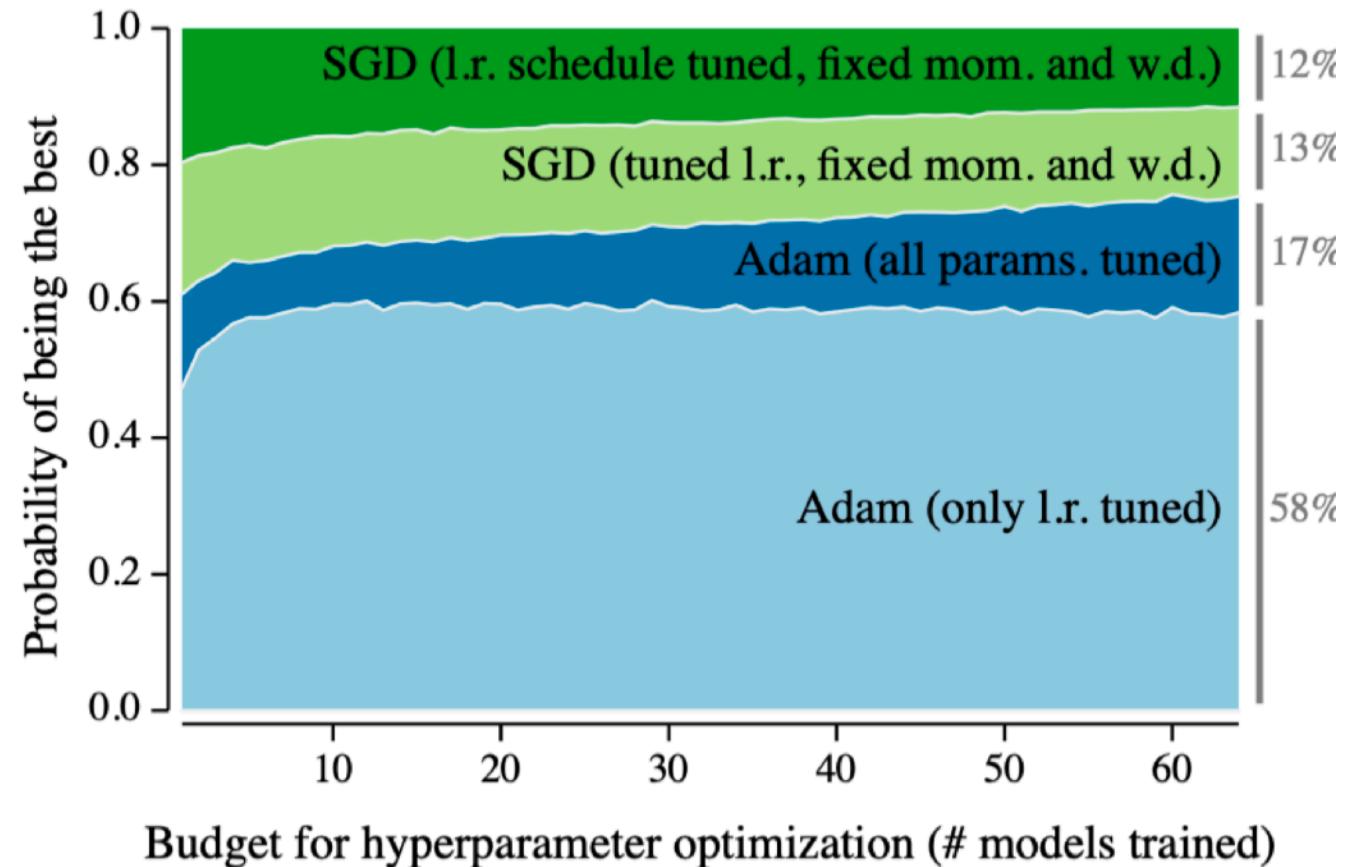
## ENCODER



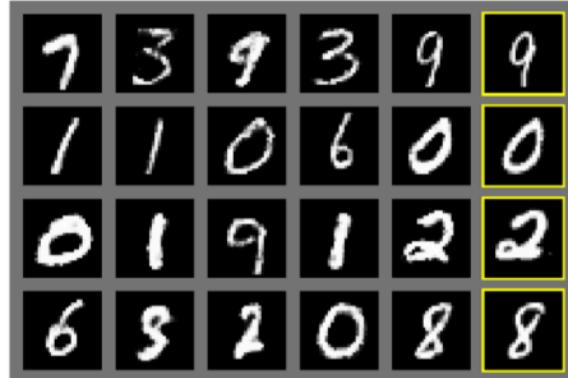
## DECODER

# 2014 - Adam Optimizer

---



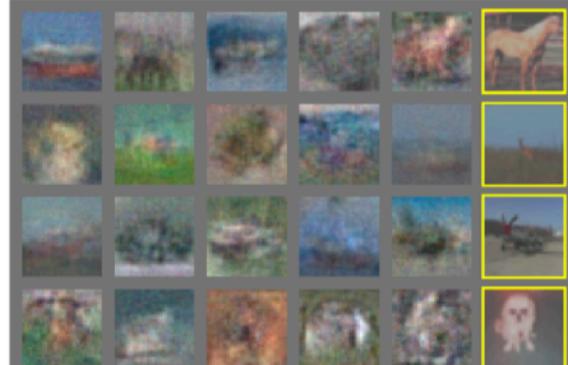
# 2015 - Generative Adversarial Network \*



a)



b)



c)

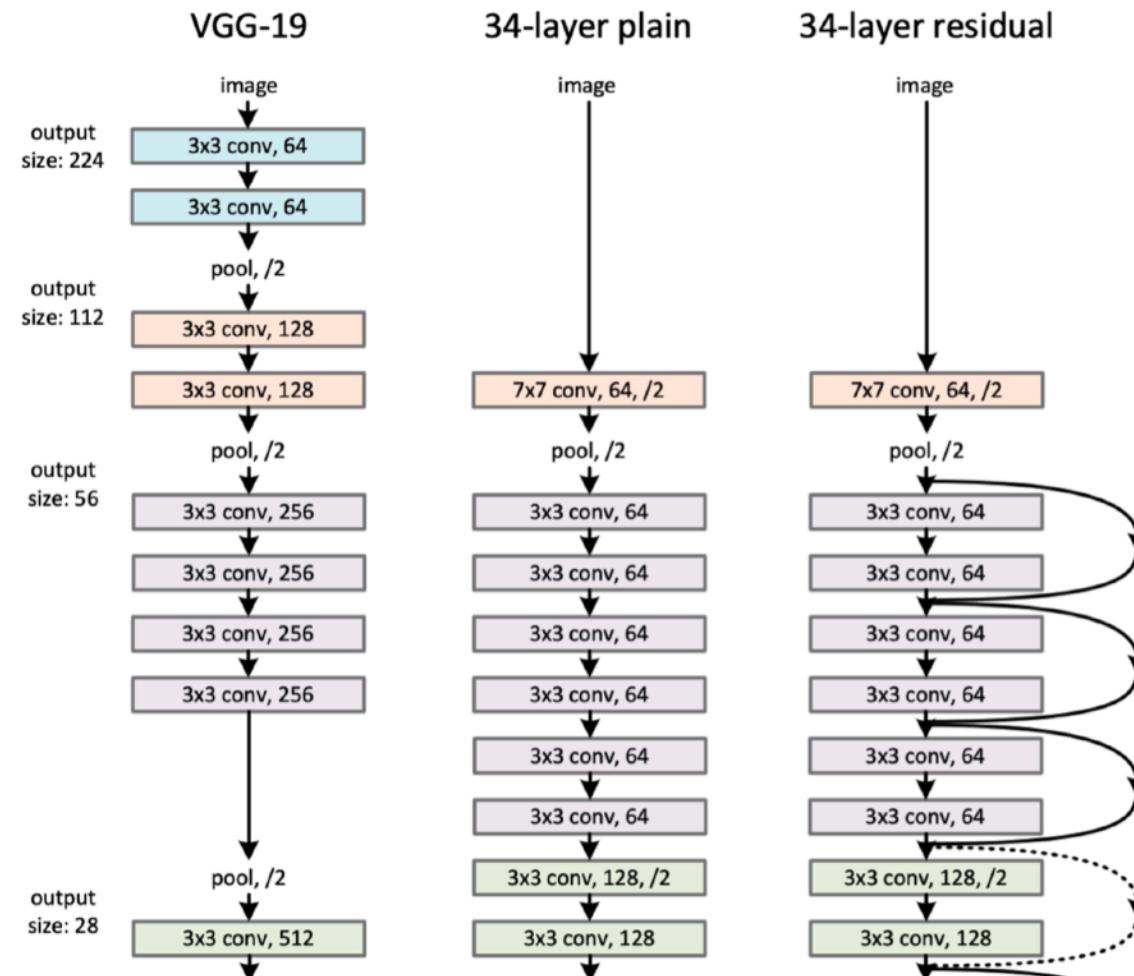


d)



“Finally, we would like to thank **Les Trois Brasseurs** for stimulating our creativity.”

# 2015 - Residual Networks



# 2017 - Transformer

## 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

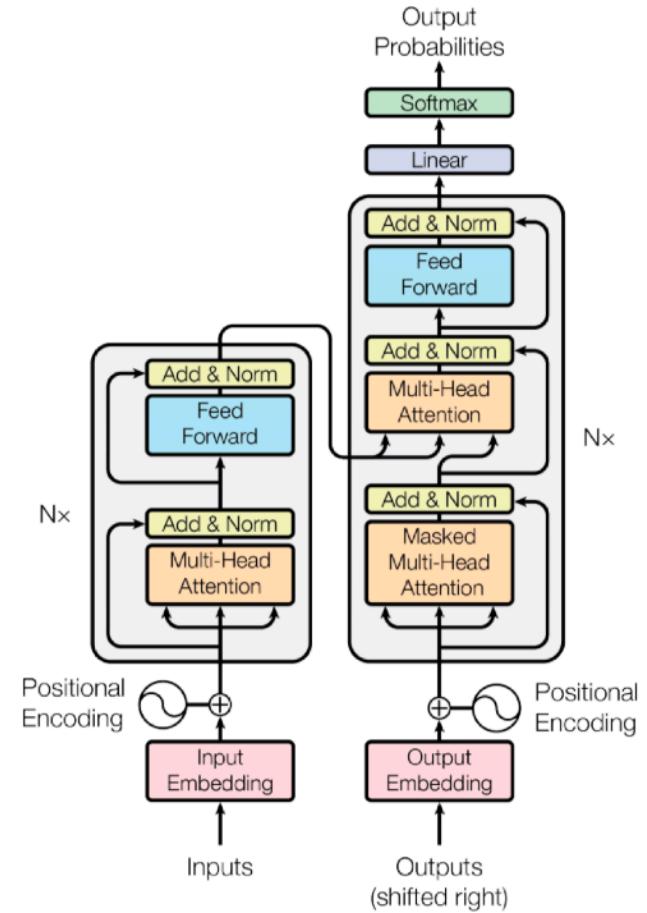
**Jakob Uszkoreit\***  
Google Research  
usz@google.com

**Llion Jones\***  
Google Research  
llion@google.com

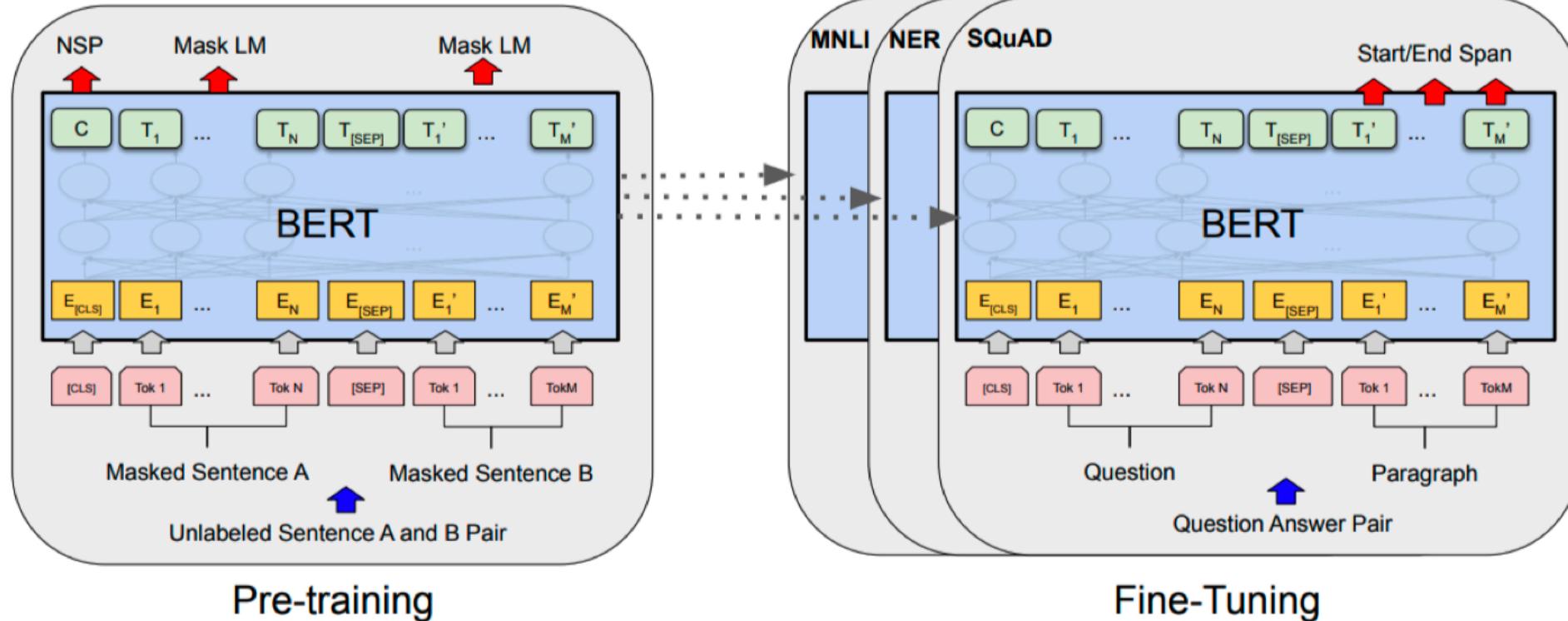
**Aidan N. Gomez\* †**  
University of Toronto  
aidan@cs.toronto.edu

**Lukasz Kaiser\***  
Google Brain  
lukaszkaiser@google.com

**Illia Polosukhin\* ‡**  
illia.polosukhin@gmail.com



# 2018 - BERT (fine-tuned NLP models)



**B**idirectional **E**ncoder **R**epresentations from **T**ransformers

# 2019 - BIG Language Models

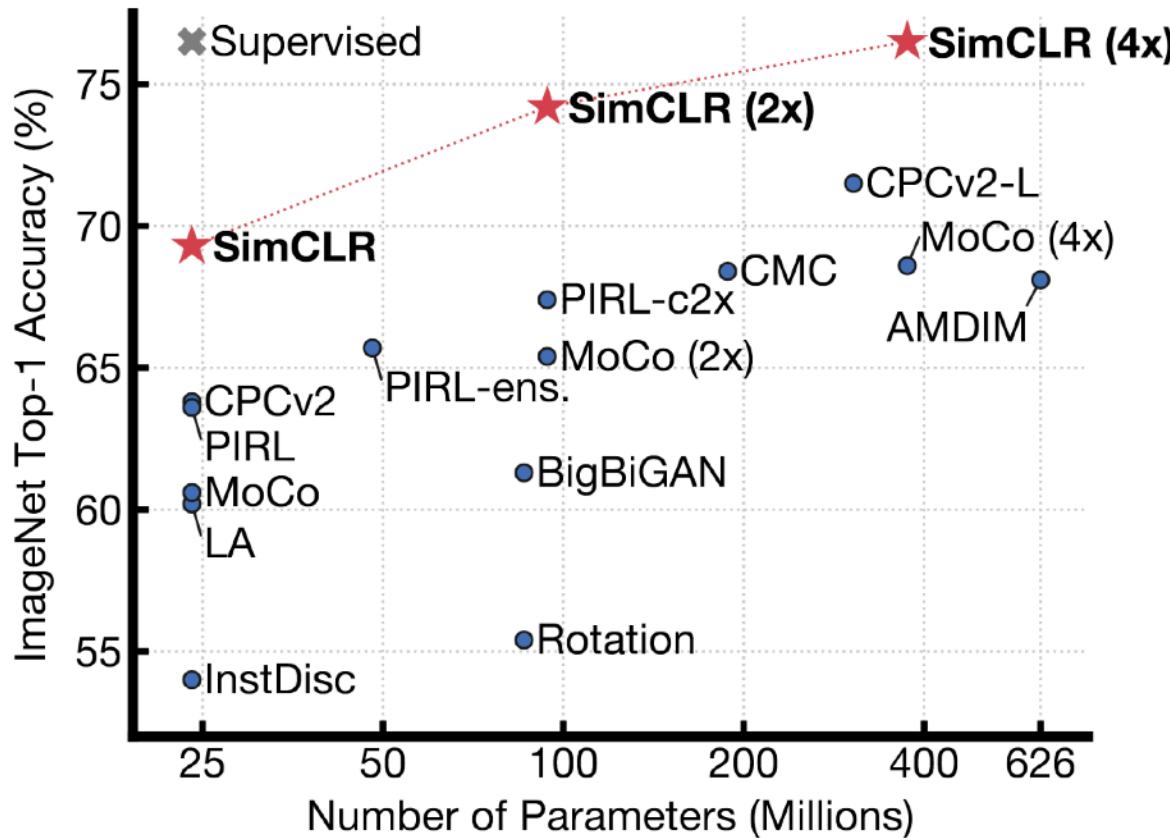
---



# OpenAI

**GPT-3, an autoregressive language model with 175 billion parameters**

# 2020 - Self Supervised Learning



**SimCLR**: a simple framework for contrastive learning of visual representations

# Thank you for listening

---