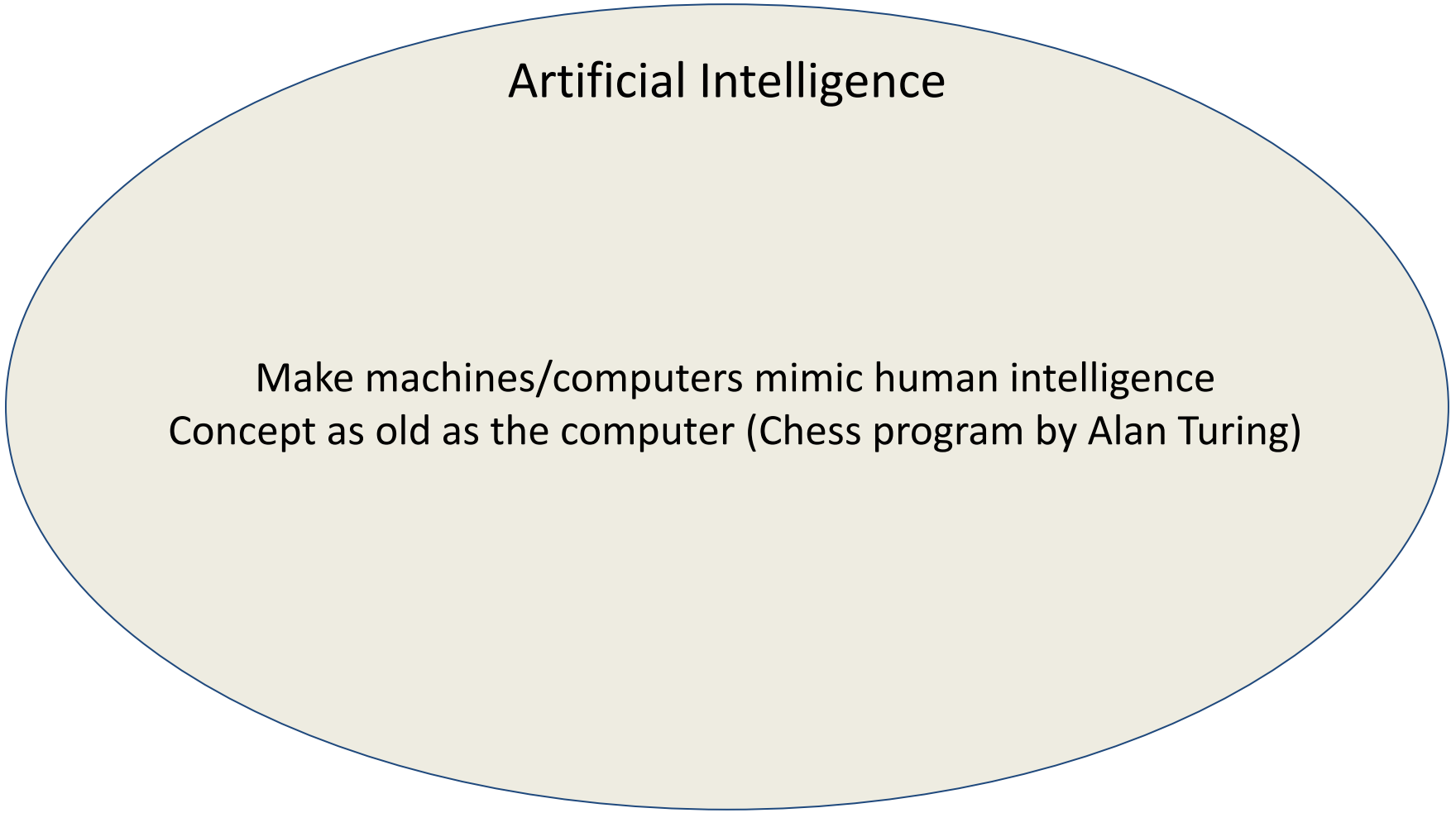
AI504: Programming for Artificial Intelligence

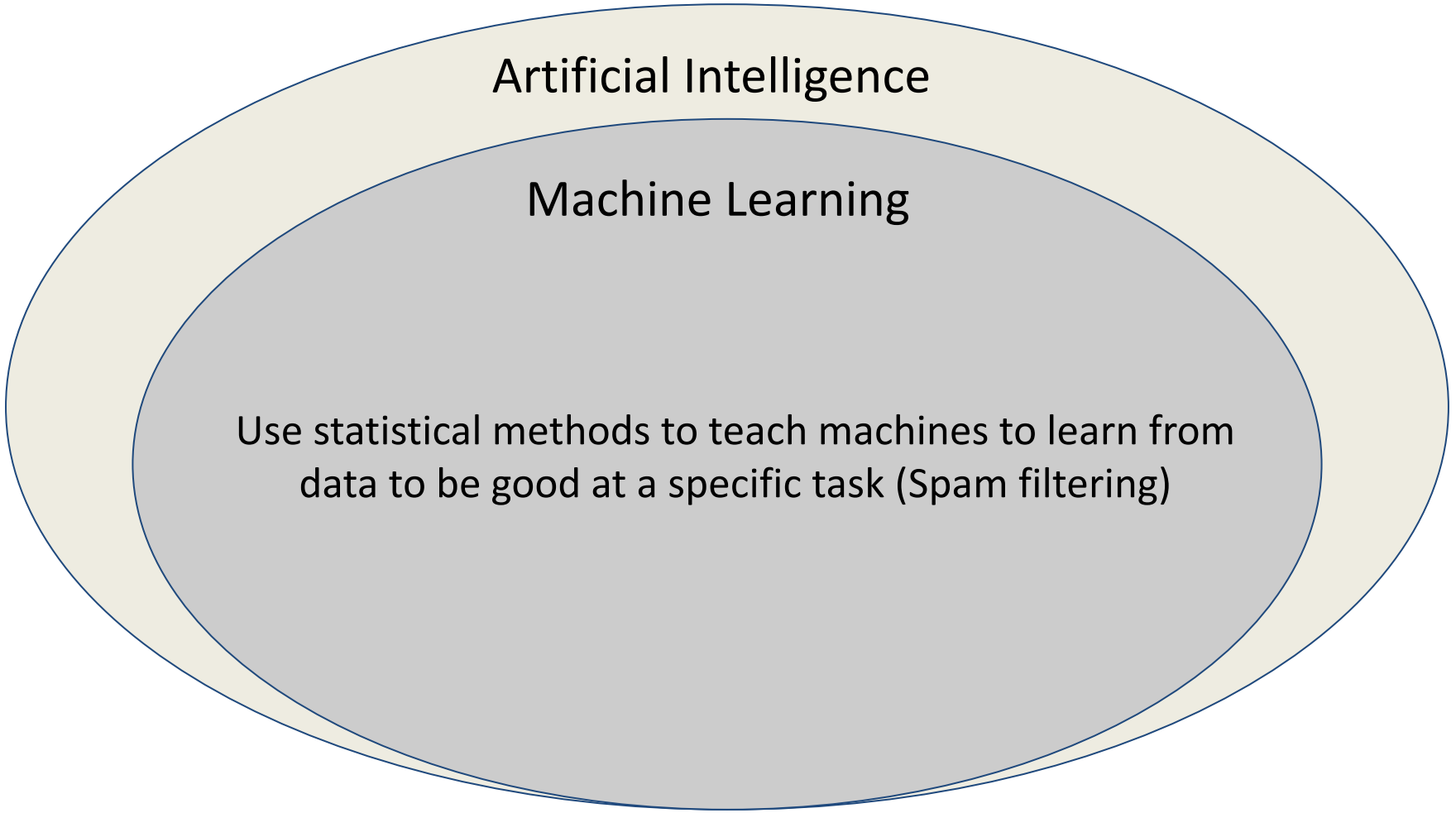
Week 1: Introduction

Edward Choi
Grad School of AI
edwardchoi@kaist.ac.kr

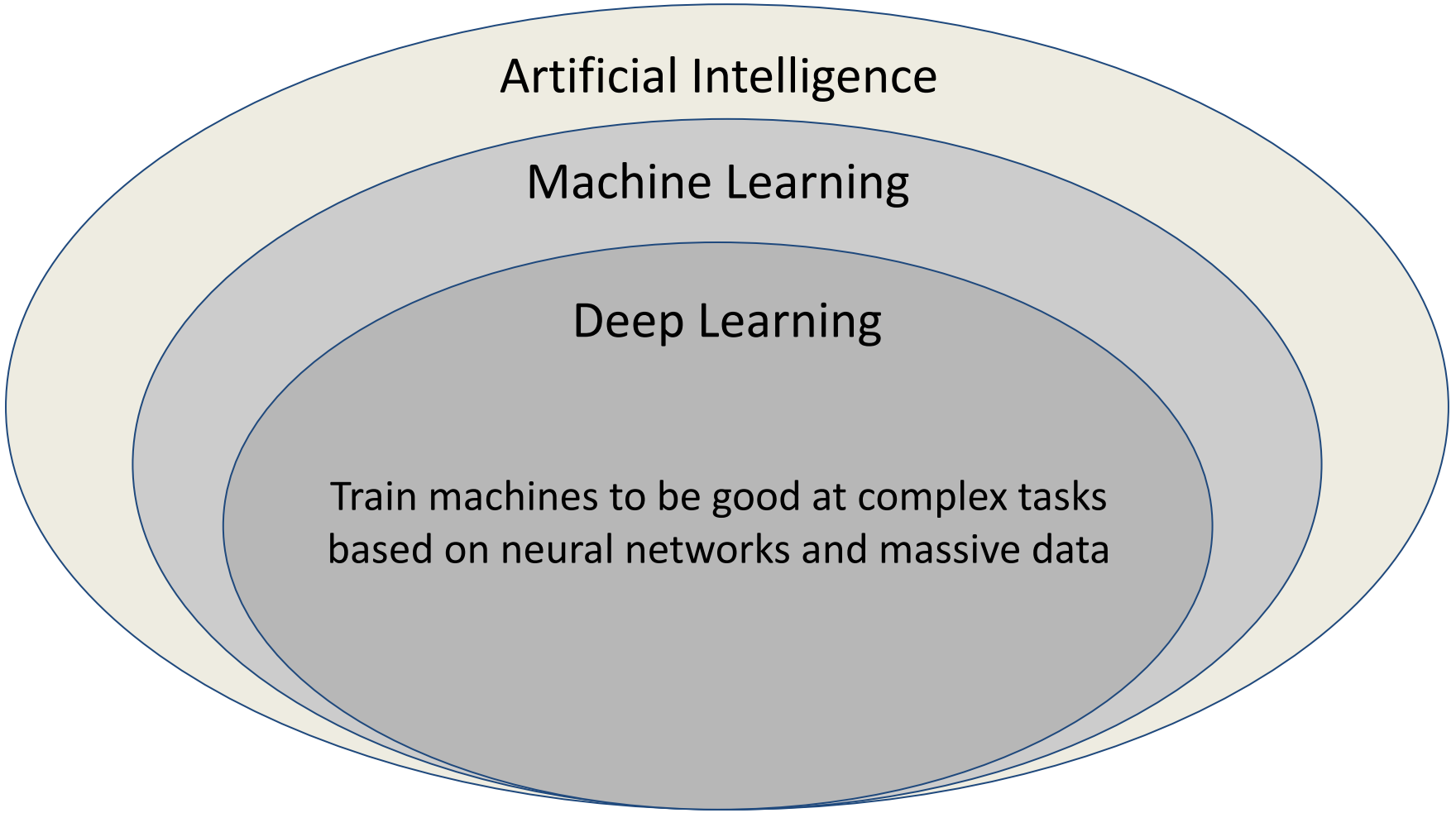
What is AI?


Artificial Intelligence
Make machines/computers mimic human intelligence
Concept as old as the computer (Chess program by Alan Turing)

What is Machine Learning?

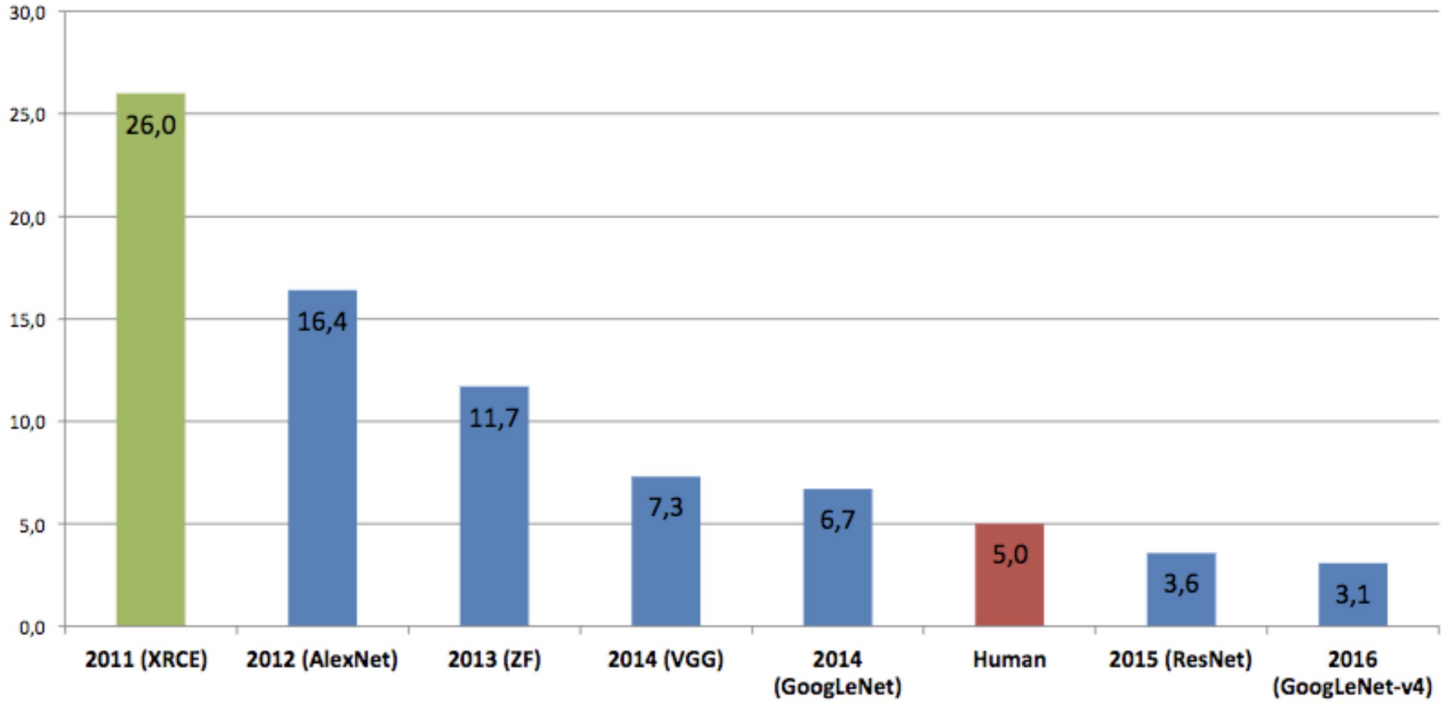


What is Deep Learning?



Why Deep Learning?

State-of-the-art performance



ImageNet Classification Error

Survey of neural networks in autonomous driving, Gustav von Zitzewitz,
Advanced Seminar SS 2017: Survey of Neural Networks in Autonomous Driving

Why Deep Learning?

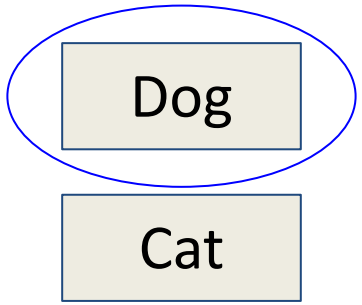
Less feature enginnering

Input Feature Extraction Classification

Classical machine learning process

Why Deep Learning?

Less feature engineering



Input

Feature Extraction + Classification

Deep learning process

How is this Possible?

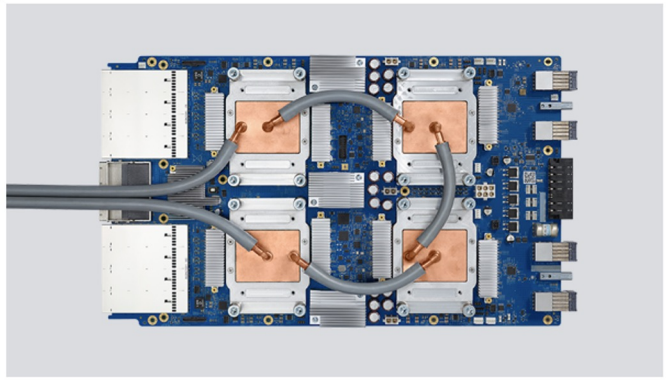
Large data

- Social network services
  - Youtube, Instagram, Twitter
- Collective intelligence
  - Wikipedia
- Mass media
  - News articls

ImageNet

How is this Possible?

Large data + Powerful machines



Cloud TPU v3

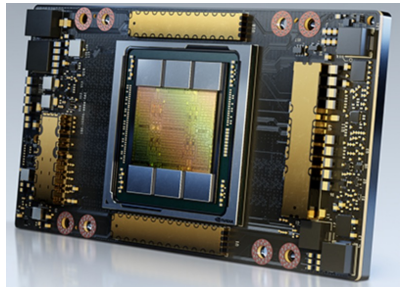
420 teraflops

128 GB HBM

NVIDIA A100

★★★★★ (2)

Details | Reviews | Stores

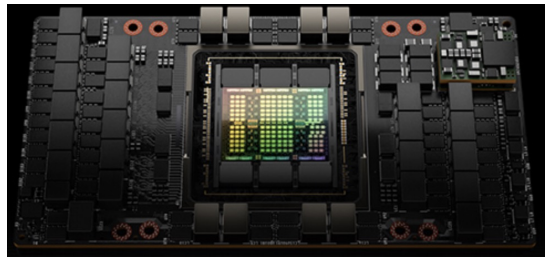


Typically $14,575–$16,041

NVIDIA H100

★★★★★ (2)

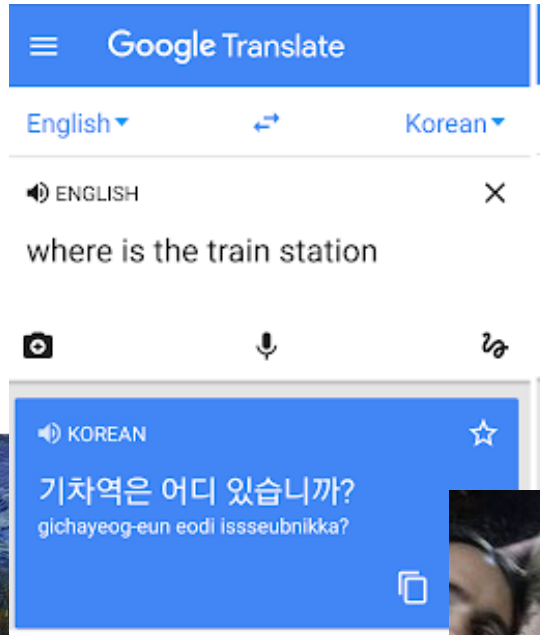
Details | Reviews | Stores

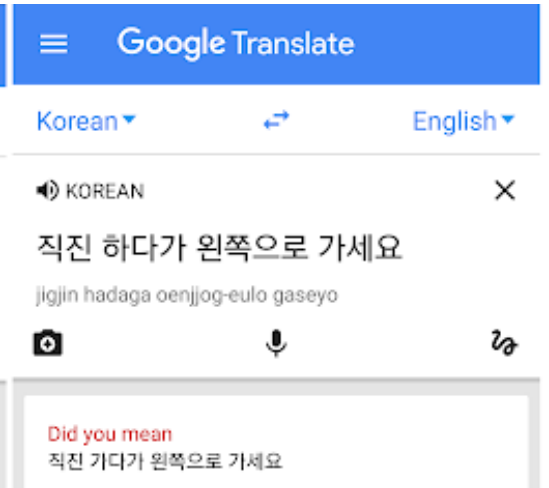


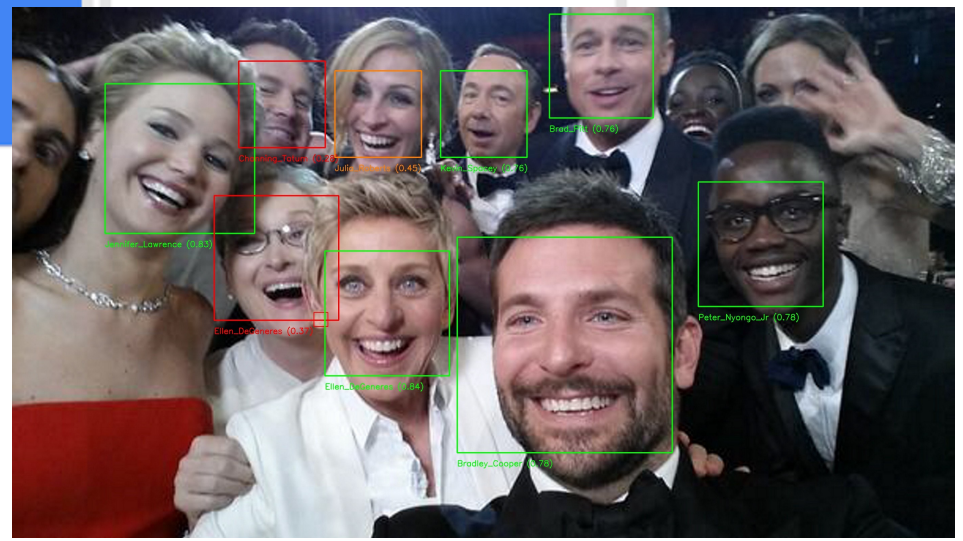
Typically $30,001–$39,000

Modern AI


ALPHAGO
2017
BFI LONDON FILM FESTIVAL
WINNER
MILL VALLEY FILM FESTIVAL
Official Selection
VIFF

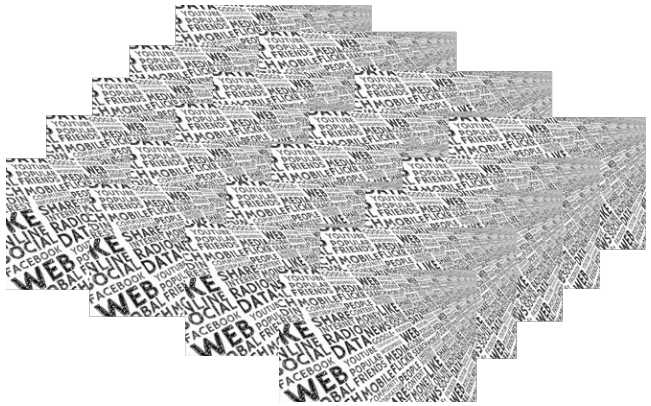

Google Translate
English
Korean
ENGLISH
where is the train station
KOREAN
기차역은 어디 있습니까?
gichayeog-eun eodi issseubnikka?


Google Translate
Korean
English
KOREAN
직진 하다가 왼쪽으로 가세요
jigjin hadaga oenjjog-eulo gaseyo
Did you mean
직진 기다가 왼쪽으로 가세요

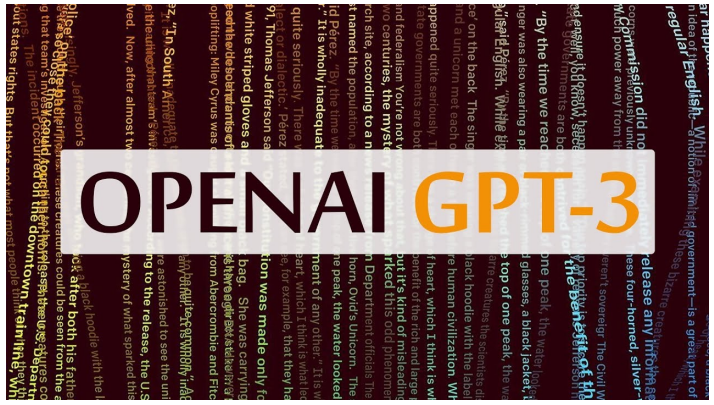


(Almost) Infinite Compute + (Almost) Infinite Data

Probably >1000 V100's



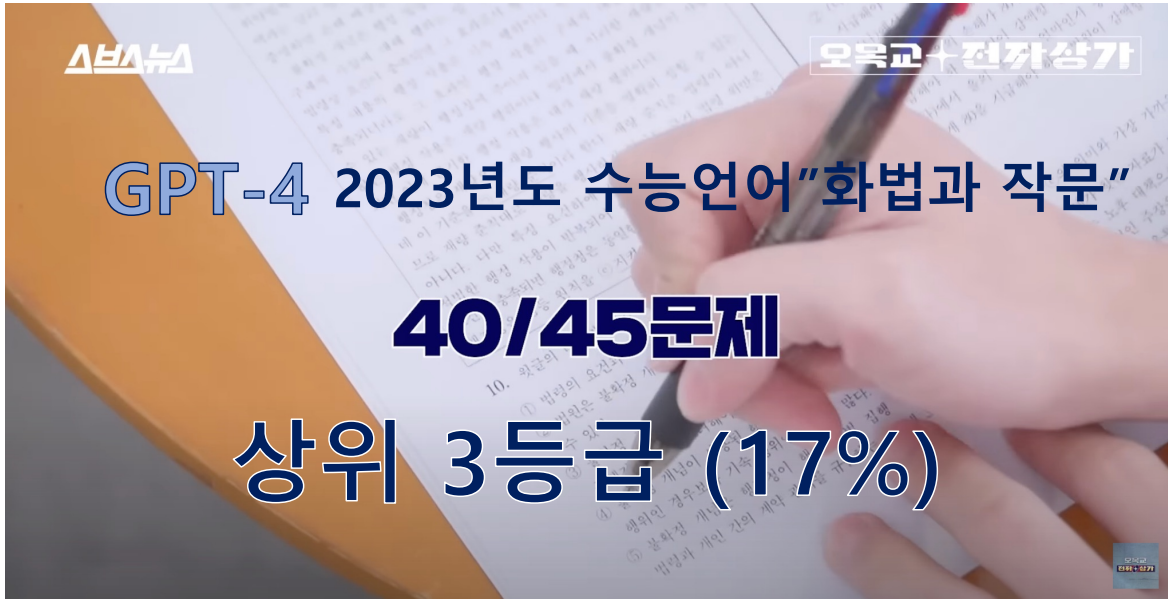
570GB Text (400 Billion Tokens)

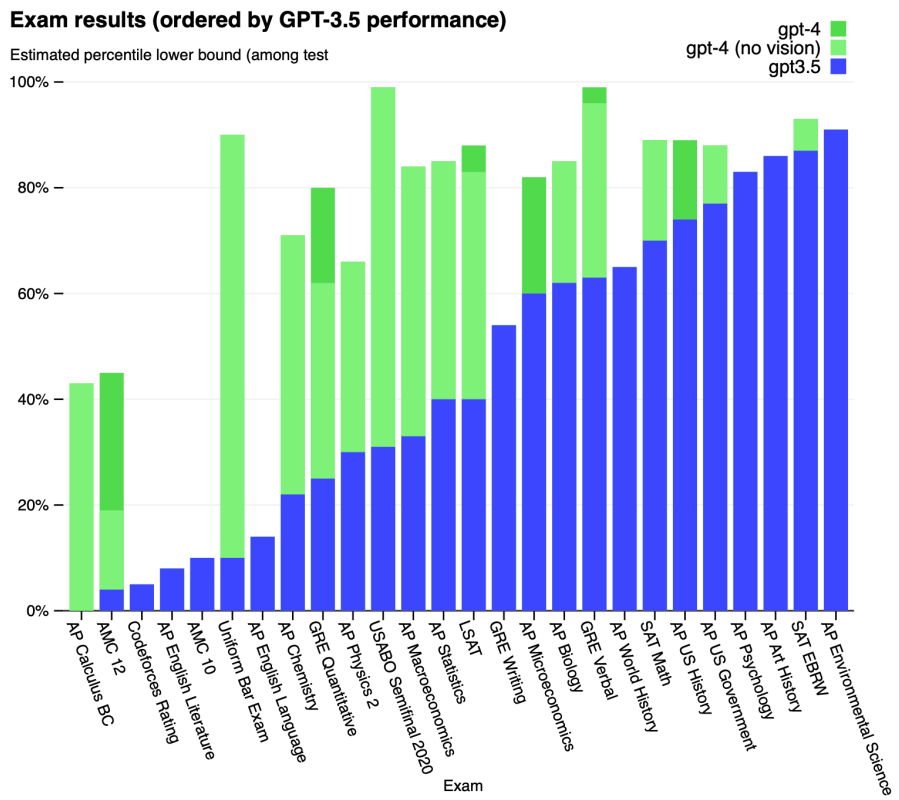


175B Parameters
$4~5 Million for Training

Large Language Models (LLMs)

- The most powerful (near) Artificial General Intelligence
  - Using tens (or even hundreds) of layers of Transformer
  - Param size 7B ~ 560B





DALL-E 2, Imagen, Stable Diffusion

- Conditional text-to-image generator
  - Diffusion + Transformer

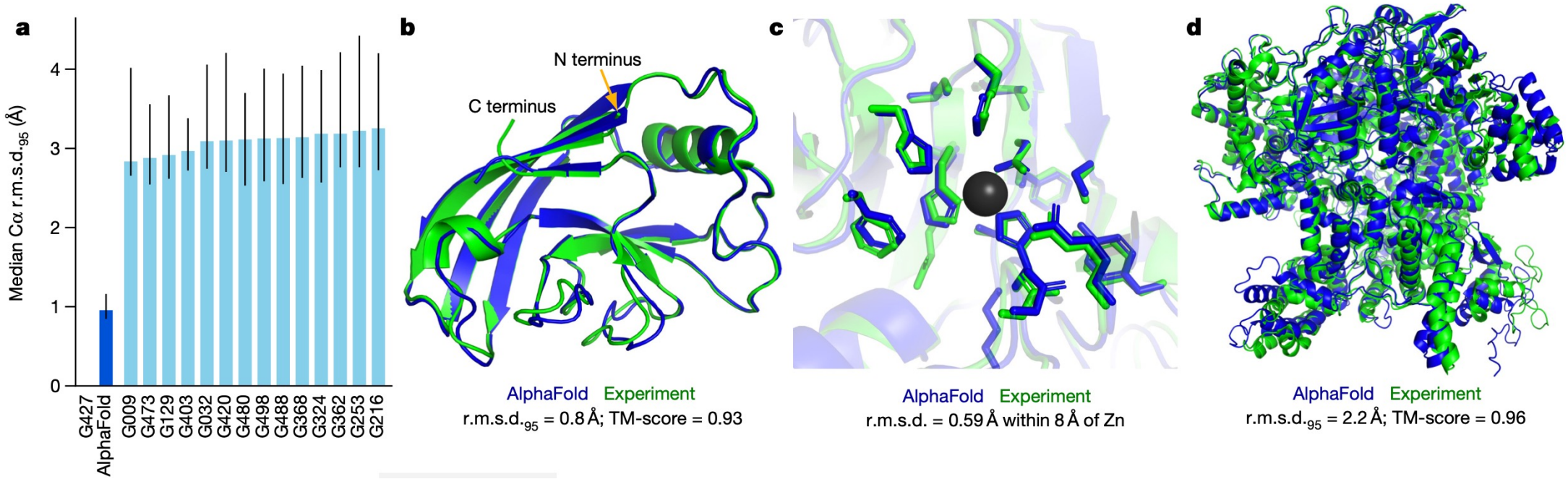
An astronaut riding a horse in a photorealistic style

A chrome-plated duck with a golden beak arguing with a golden turtle in a forest

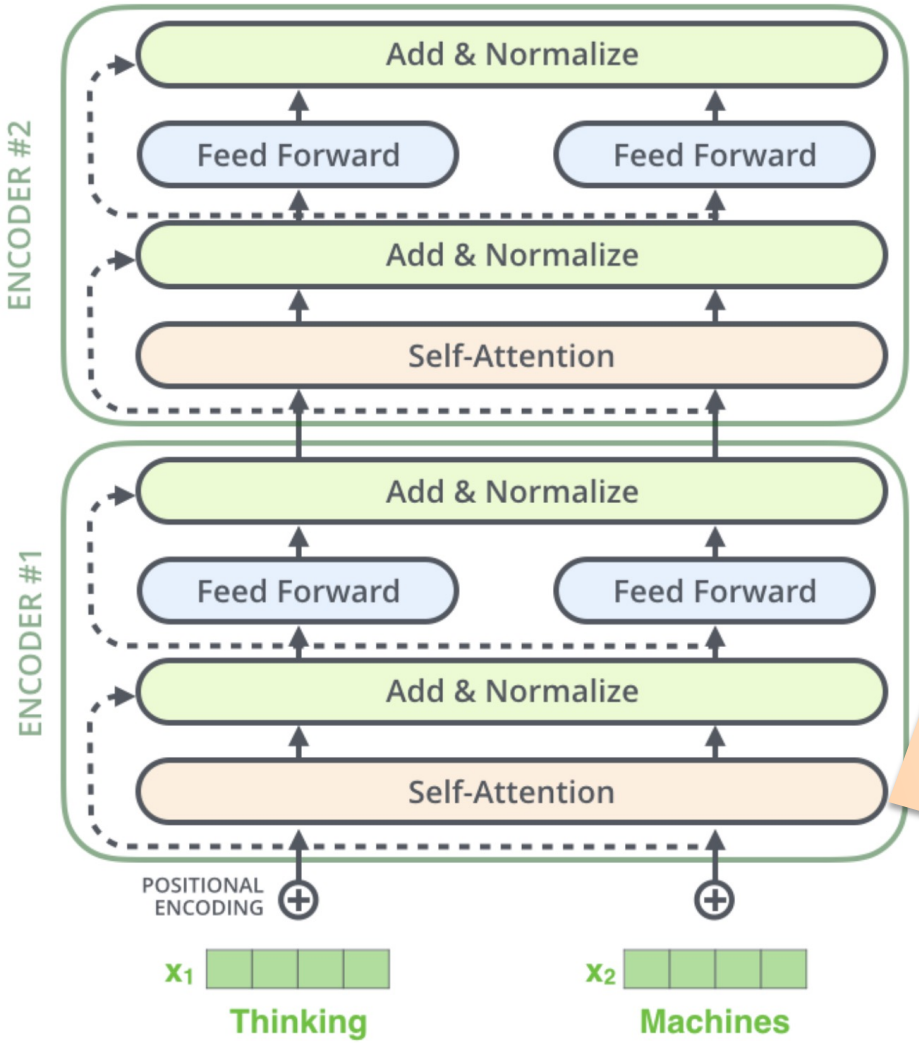
A photograph of an astronaut riding a horse

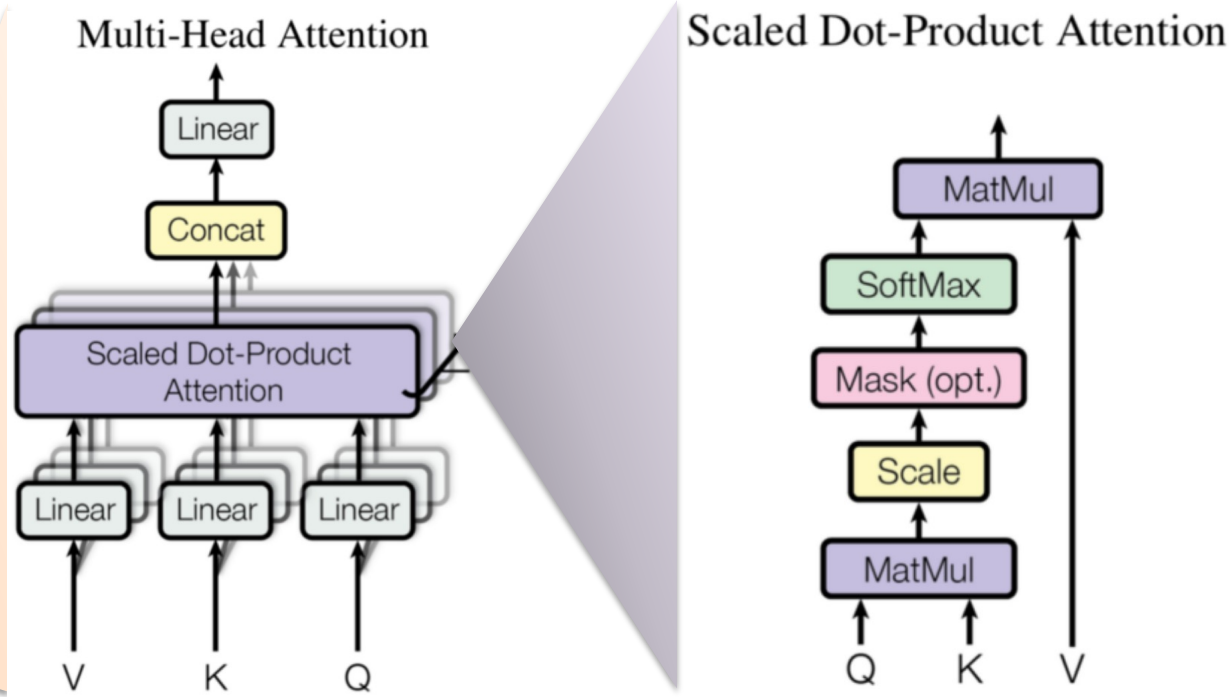
AlphaFold2

- DeepMind's protein structure prediction model
  - Using 48 layers of Evoformer
    - Special self-attention for graphs in 3D



Transformer Architecture





http://jalammar.github.io/illustrated-transformer/

Goal

- Learn to build deep learning models.
  - So that you can replicate papers
  - So that you can realize your ideas
  - So that you can conduct AI research

- This course teaches only the very basics.
  - Practice makes perfect!

Structure

- Schedule
  - Lecture on Tuesday
    - Conducted by the lecturer
  - Practice on Thursday
    - Conducted by the TA
- Evaluation
  - Letter grade
    - Project 1: 30%, Project 2: 30%, Project 3: 40%
  - No attendance check
    - Up to you to attend
- Projects
  1. Image synthesis using GAN
  2. Language model using Transformer
  3. Visual Large Language Model (Visual LLM)
  - All projects are given 24hr
  - Expected to take 2~3hr (training will take ~5min on Colab)
  - Designed to evaluate only model programming, not model performance

Weekly Plan

1. Intro + Numpy
2. Basic Machine Learning + Scikit-learn
3. PyTorch Intro + Logistic Regression + Multi-layer Perceptron
4. Autoencoders (& Denoising Autoencoders)
5. Variational Autoencoders
6. Generative Adversarial Networks
7. Convolutional Neural Networks
8. **Project 1: Image Synthesis**
9. Word2Vec + Subword Encoding
10. Recurrent Neural Networks & Sequence-to-Sequence
11. Transformers
12. BERT & GPT
13. **Project 2: Language Model**
14. Deep Diffusion Probabilistic Model
15. Image-Text Multimodal Learning
16. **Project 3: Visual Language Model**

Extra: Neural ODE, Graph Neural Networks ➔ Will share video only

Expectation

- Lecture attendance: Up to you
  - 360 people enrolled in this course
  - Last year material: https://mp2893.com/course.html
  - Flip learning (Edu 4.0)?

- 3 Projects
  - All projects are given 24 hours to finish
  - Individual effort
    - Cheating/copying, if caught, will lead to **serious consequences**
  - Deliverables
    - Your code
    - Your output
      - Project 1: images
      - Project 2: tokens
      - Project 3: tokens

Material & Discussion

- No textbook
- Materials will be posted on Classum if necessary.

- Classum
  - Join as Participant Link: www.classum.com/2GKAC7FYT

Teaching Team

- Lecturer
  - Edward Choi
    - edwardchoi@kaist.ac.kr
    - https://mp2893.com
- TA
  - TBA

Edward Choi, 최윤재

- Education
  - Ph.D. (computer science), Georgia Tech, 2014-2018
    - Thesis: Interpretable deep learning for longitudinal electronic health records
- Professional Experience
  - ETRI (2010-2014)
  - Sutter Health (2015, 2016)
  - DeepMind & Google (2017)
  - Google Brain & Google Health (2018-2020)
- Research Area
  - Machine Learning, Healthcare, NLP, Multi-modal

First Assignment

- Install Anaconda
  - Python package for data science
  - Includes Jupyter, Numpy, Scikit-Learn, TensorFlow, PyTorch
  - https://www.anaconda.com/products/individual

- All practice sessions will be conducted with Google Colab
  - Python Notebook on the web
  - Can train models using Google's GPU/TPU
  - Session-based (why you need Anaconda)

Questions?

AI504: Programming for Artificial Intelligence

Week 1: Introduction

Edward Choi
Grad School of AI
edwardchoi@kaist.ac.kr