AI-504

Introduction

Week 01

인하대학교 통계학과 김현수

스터디 목표 : AI504 + Project
1강 : Introduction
2강 : Numpy
3~4강 : Sklearn + Practice => 1주차

5~6강 : Pytorch - Logreg, NN + Practice
7~8강 : AutoEncoder+Practice => 2주차

9~10강 : Variational AutoEncoder + Practice
11~12강 : GAN + Practice => 3주차

13~14강 : CNN + Practice
15~16강 : Word Embedding + Practice => 4주차

17~18강 : RNN + Practice
19~20강 : Img2Txt + Practice => 5주차

21~22강 : Transformer + Practice
23~24강 : BERT & GPT + Practice => 6주차

25~26강 : Graph NN + Practice
27~28강 : Neural ODE + Practice => 7주차

이후 AI Hub에 있는 데이터로 자율 프로젝트 진행 => 8~9주차

Contents

- AI vs ML vs DL (vs RL)

- 딥러닝을 배워야 하는

이유

- 딥러닝의 응용분야

- 딥러닝 프레임워크

 # AI vs ML vs DL (vs RL)

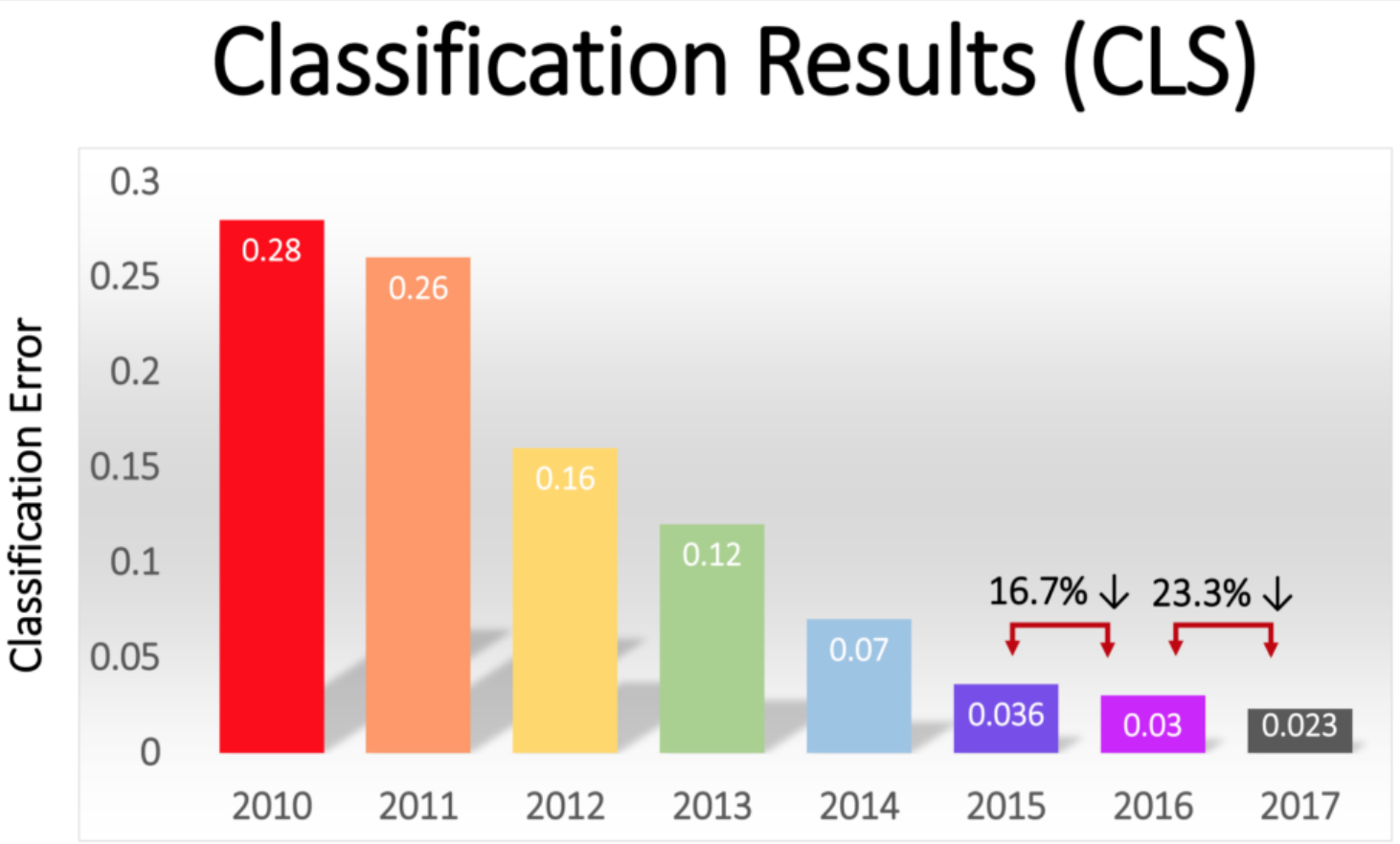
Differe

	Deep Learning	Reinforcement Learning
공통점	• Autonomous, Self-Teaching System	
차이점	• Training Set으로부터 학습 • 학습을 새로운 데이터에 적용	• 최고의 보상을 위한 행동 선택 • 동적으로 학습하며 행동을 조정

딥러닝을 배워야 하는 이유

Why
DEEP Learning

2015 ResNet
-> 8360만 변수, 80% 정확도
2021 CoAtNet
-> 14억~24억 변수, 90~91% 정확도

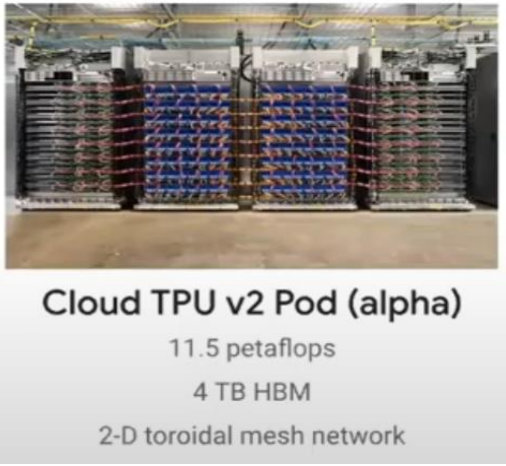


참고:https://wikidocs.net/147236

How is this possible?

Large Data + Powerful Machines

Large Data

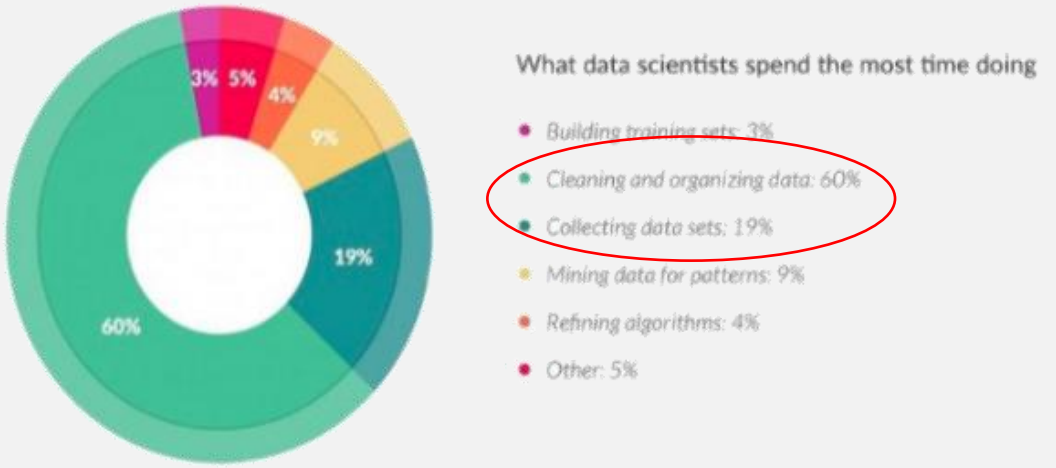


Powerful Machines

 # "garbage in, garbage out!"

Why DEEP Learning

데이터 사이언티스트들은
데이터 준비에
약 80%의 시간을
투자한다.



Source: https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/

참고 :
https://velog.io/@guide333/%EC%95%84%EC%9D%B4%ED%9A%A8-Feature-Engineering

 # "garbage in, garbage out!"

Why DEEP Learning

The techniques of Feature Engineering

1. Imputation
2. Outliers
3. Binning
4. Log Transform
5. One-hot Encoding
6. Grouping Operagion
7. Feature Split
8. Scaling
9. Extracting Data

Feature Engineering

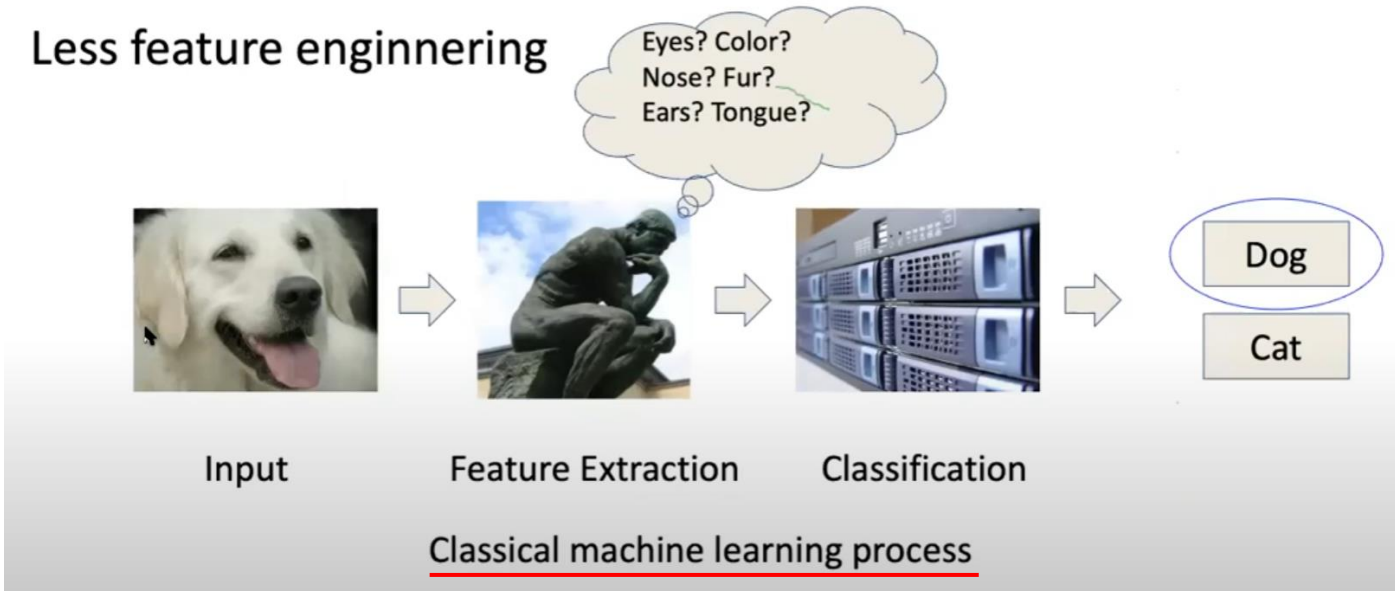
모델 정확도를 높이기 위해서 주어진 데이터를
예측 모델의 문제를 잘 표현할 수 있는 features로 변형시키는 과정

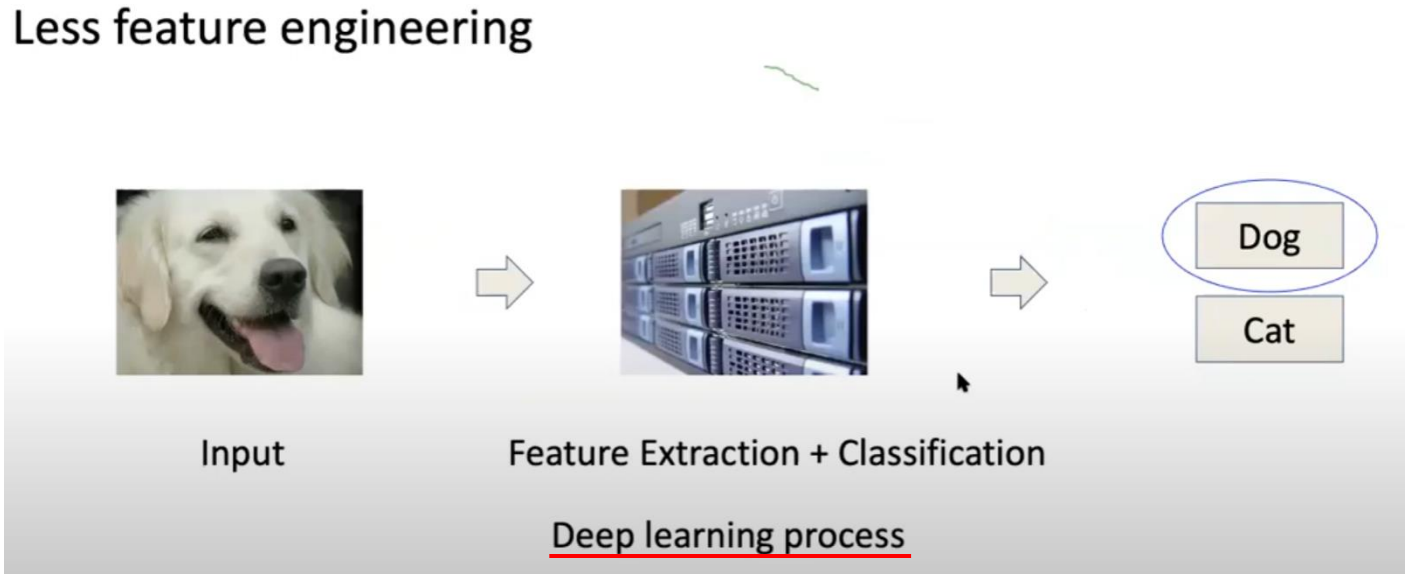
Goals of Feature Engineering

1. 머신 러닝 알고리즘에 걸맞는 적당한 입력 데이터셋을 준비
2. 머신 러닝 모델의 성능을 향상시키는 것

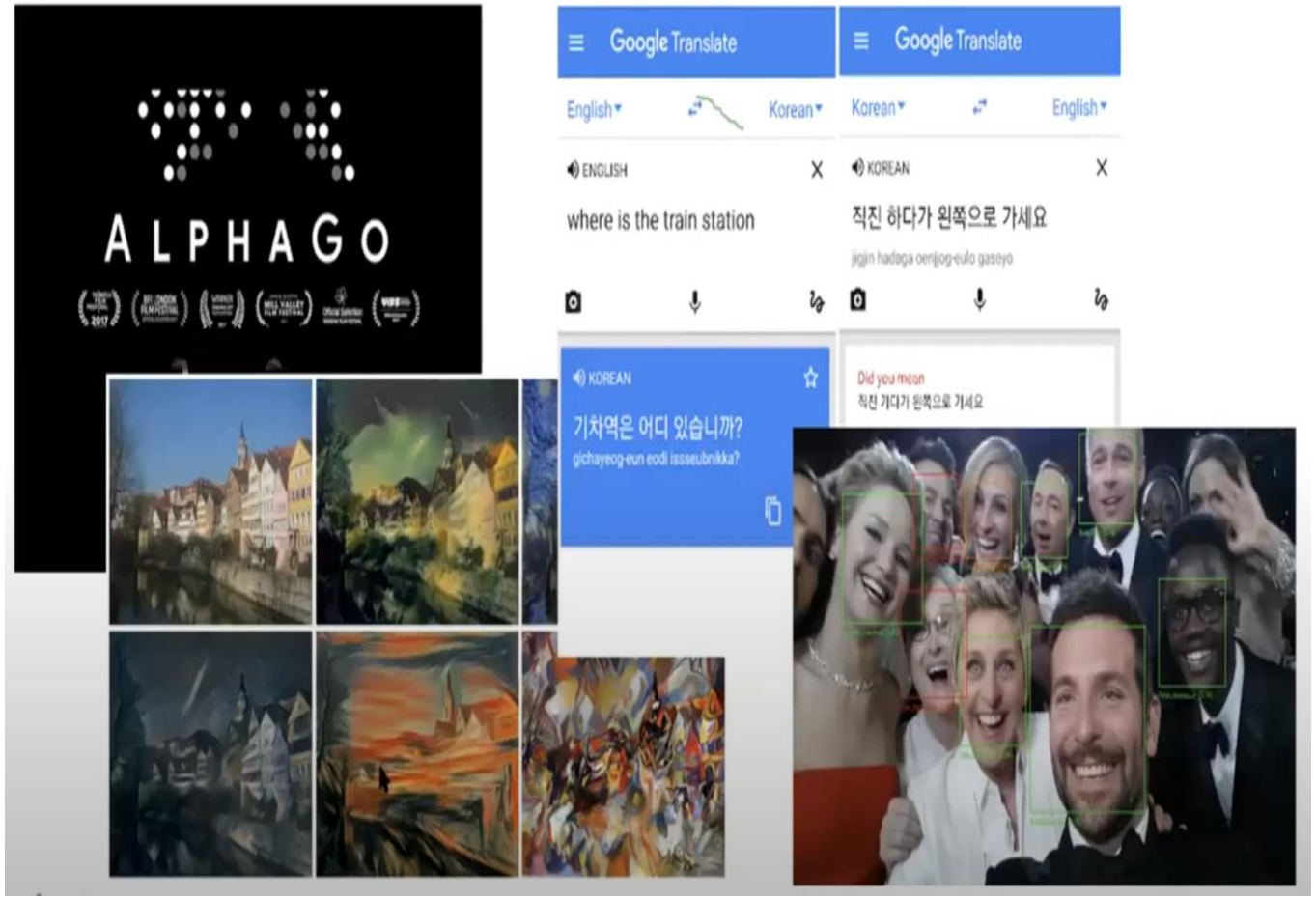
 # "garbage in, garbage out!"

Why DEEP Learning



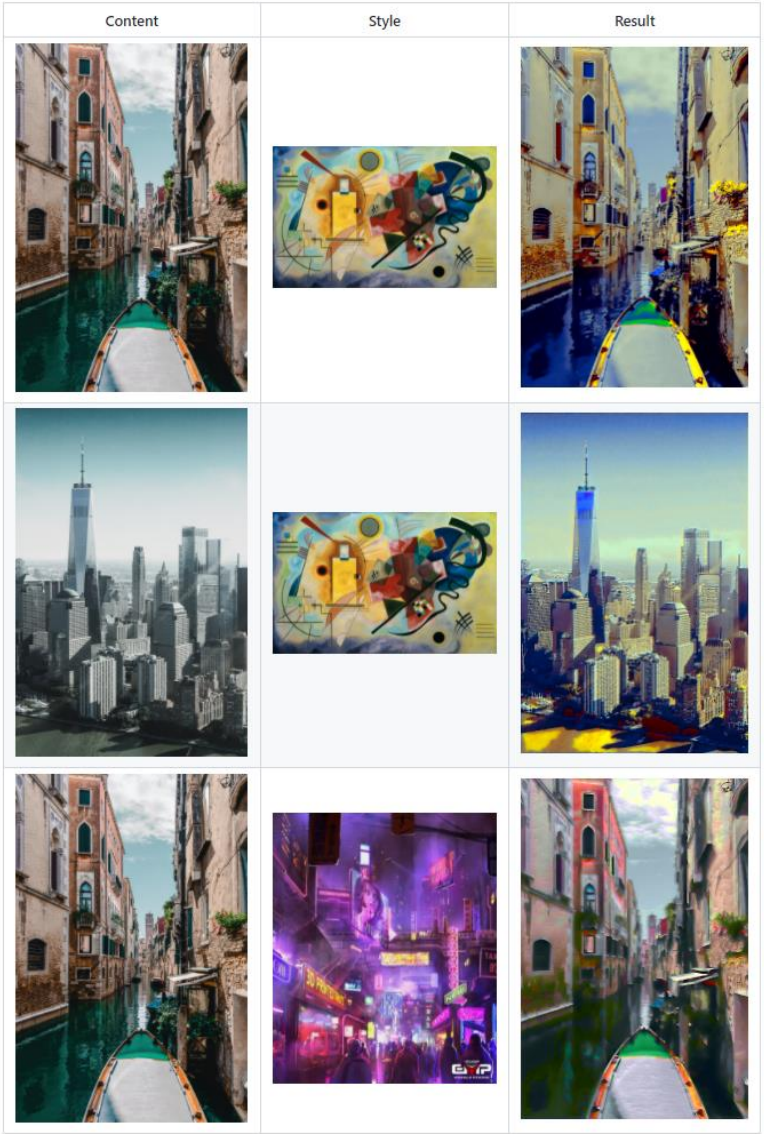


Modern AI



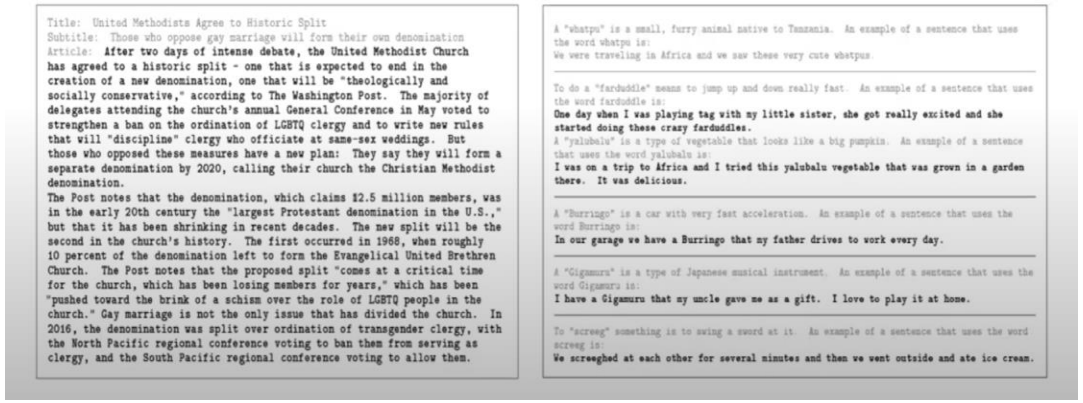
https://github.com/kairess/style-transfer

The simplest code for neural style transfer code using PyTorch.



GPT3

- Open AI's unsupervised generative language model
  - Using 96 layers of Transformer
  - BERT uses 24 layers



Title: United Methodists Agree to Historic Split
Subtitle: Those who oppose gay marriage will form their own denomination
Article: **After two days of intense debate, the United Methodist Church has agreed to a historic split - one that is expected to end in the creation of a new denomination, one that will be "theologically and socially conservative," according to The Washington Post. The majority of delegates attending the church's annual General Conference in May voted to strengthen a ban on the ordination of LGBTQ clergy and to write new rules that will "discipline" clergy who officiate at same-sex weddings. But those who opposed these measures have a new plan: They say they will form a separate denomination by 2020, calling their church the Christian Methodist denomination.**
The Post notes that the denomination, which claims 12.5 million members, was in the early 20th century the "largest Protestant denomination in the U.S.," but that it has been shrinking in recent decades. The new split will be the second in the church's history. The first occurred in 1968, when roughly 10 percent of the denomination left to form the Evangelical United Brethren Church. The Post notes that the proposed split "comes at a critical time for the church, which has been losing members for years," which has been "pushed toward the brink of a schism over the role of LGBTQ people in the church." Gay marriage is not the only issue that has divided the church. In 2016, the denomination was split over ordination of transgender clergy, with the North Pacific regional conference voting to ban them from serving as clergy, and the South Pacific regional conference voting to allow them.

A "whatpu" is a small, furry animal native to Tanzania. An example of a sentence that uses the word whatpu is:
We were traveling in Africa and we saw these very cute whatpus.

To do a "farduddle" means to jump up and down really fast. An example of a sentence that uses the word farduddle is:
One day when I was playing tag with my little sister, she got really excited and she started doing these crazy farduddles.

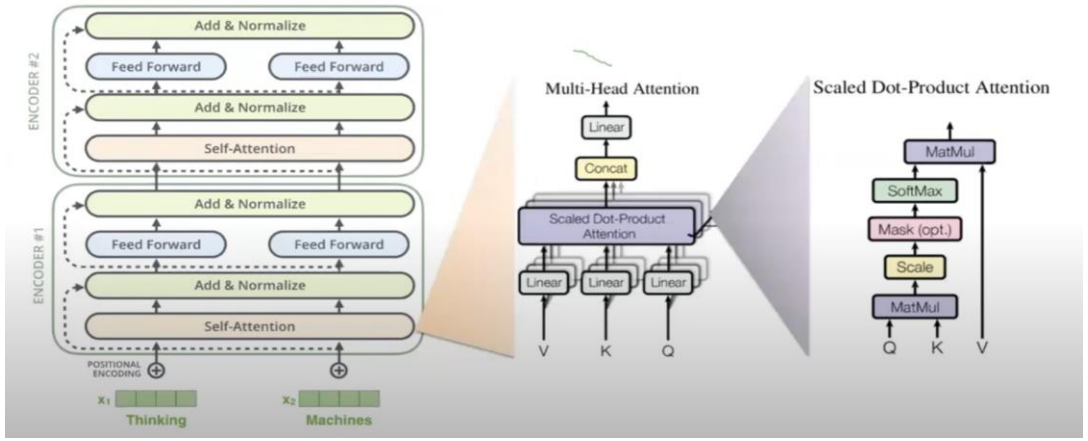
A "yalubalu" is a type of vegetable that looks like a big pumpkin. An example of a sentence that uses the word yalubalu is:
I was on a trip to Africa and I tried this yalubalu vegetable that was grown in a garden there. It was delicious.

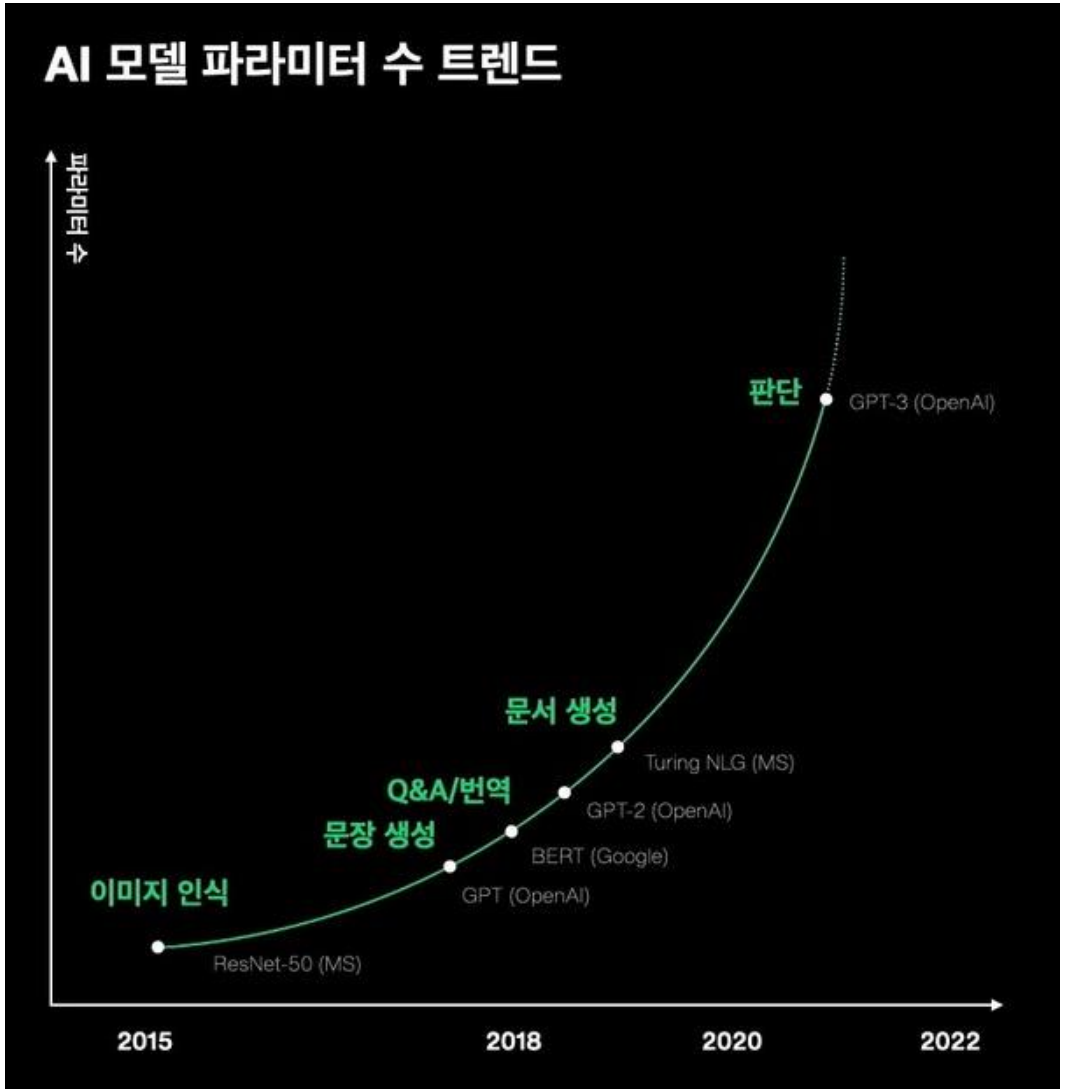
A "Burringo" is a car with very fast acceleration. An example of a sentence that uses the word Burringo is:
In our garage we have a Burringo that my father drives to work every day.

A "Gigamuru" is a type of Japanese musical instrument. An example of a sentence that uses the word Gigamuru is:
I have a Gigamuru that my uncle gave me as a gift. I love to play it at home.

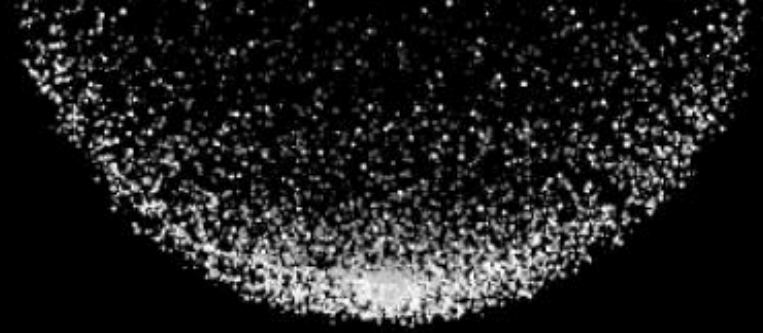
To "screeg" something is to swing a sword at it. An example of a sentence that uses the word screeg is:
We screeghed at each other for several minutes and then we went outside and ate ice cream.

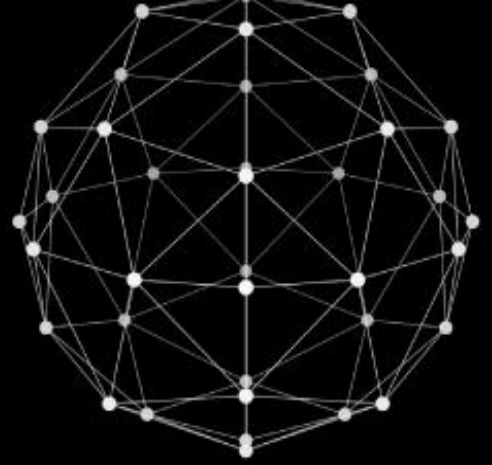
Transformer Architecture





010





HyperCLOVA

AI, 모두의 능력이 되다

- 가장 인기있는 딥러닝 라이브러리
- 2015년 구글에서 개발해 공개됨
- C++ , R, js 등 다른 언어도 지원

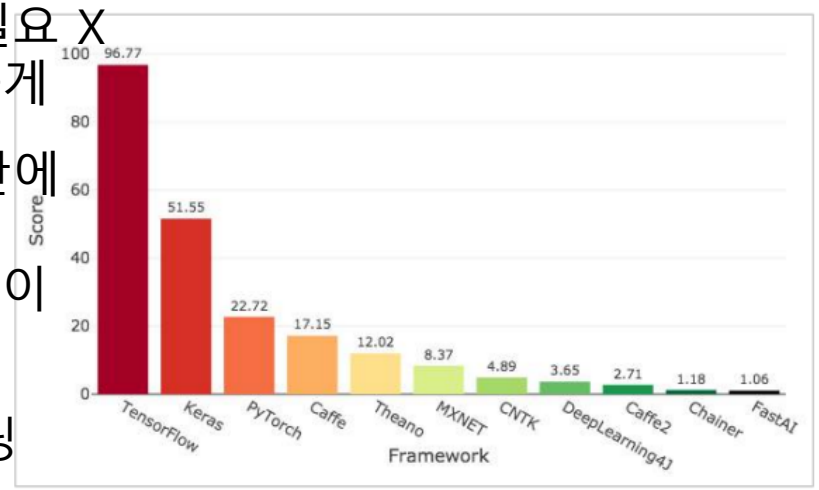
- 파이썬 기반으로 작성된 매우 가볍고 배우기 쉬운 라이브러리
- 직관적인 API로 복잡한 내부 엔진을 알 필요 X
- 너무 의존한다 보면 텐서플로우를 완전하게 익히지 못한다

최초의 딥러닝 라이브러리
- 파이썬 기반으로 CPU 및 GPU의 수치계산에 유용
- 확장성이 뛰어나지 않으며 다중 GPU지원이 부족함

- Lua 기반의 딥러닝 프레임워크인 Torch의 파이썬 버전
- 빠르고 유연한 실험을 하게 해주는 딥러닝 프레임워크

AWS에서 선택한 딥러닝 엔진
- 페이스북 AI 연구팀에 채택
- 확장성이 좋고 다중 GPU와 컴퓨터로 작업할 수 있다 -> 기업용으로 유용
- 아파치 소프트웨어 재단에서 개발

- MS에서 번역기술, 음성인식, 이미지 인식 등과 관련한 트레이닝 할때 이용
- 텐서플로, 테아노와 비교할 때 높은 확장성과 성능을 제공



[그림 1] Deep Learning Framework Power Scores(※ 출처: Towards Data Science)

감사합니다

#Thank You