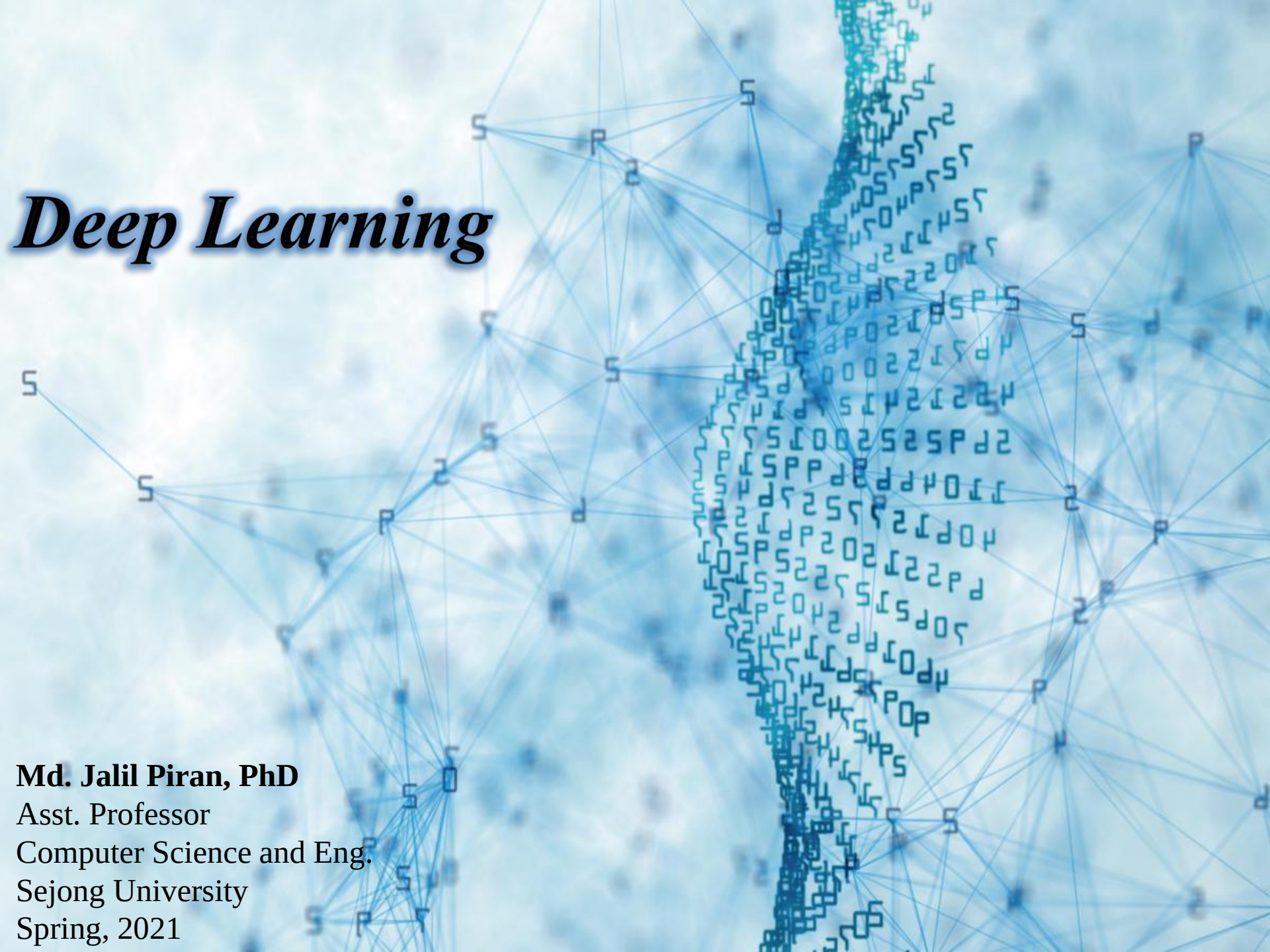


Deep Learning

Md. Jalil Piran, PhD
Asst. Professor
Computer Science and Eng.
Sejong University
Spring, 2021

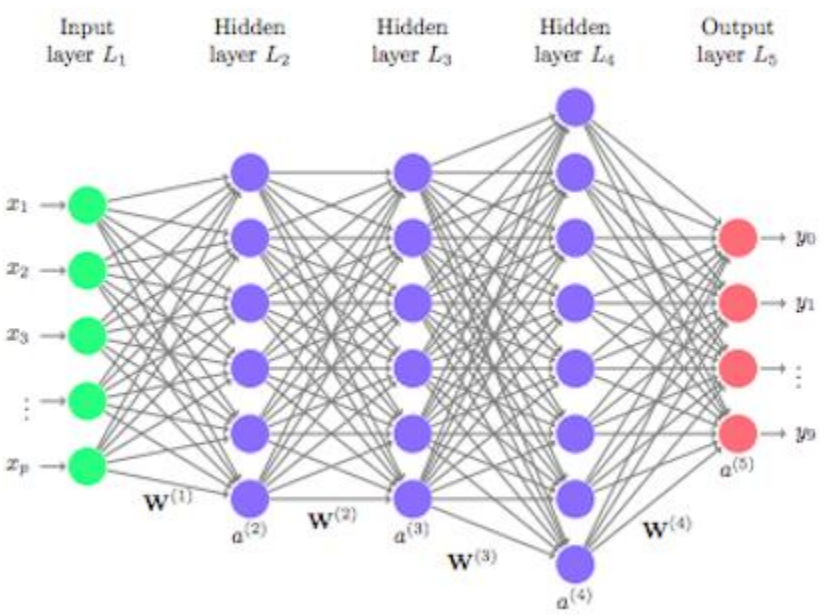
Outline

- **Introduction to Deep Learning (DL)**
- **The History of DL**

THE HISTORY OF DEEP LEARNING

Deep Learning

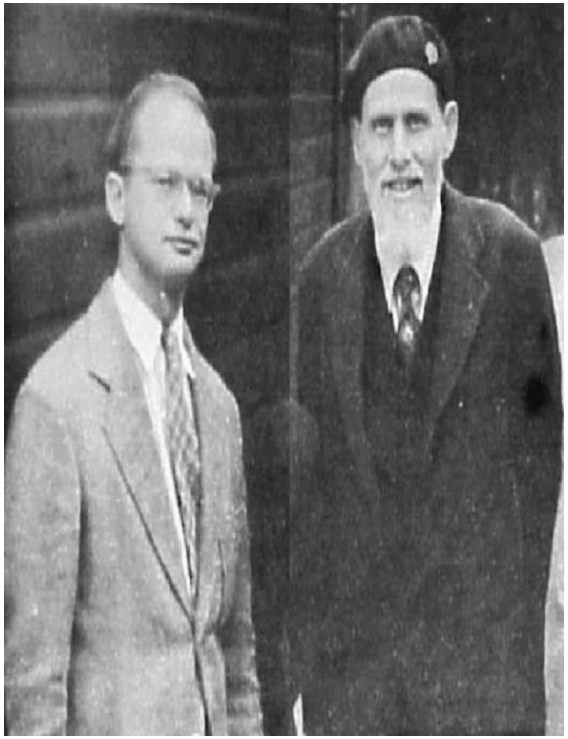
- DL is a branch of ML.
- Used to:
  - Process data
  - Imitate the thinking process
  - Understand human speech
  - Visually recognize objects

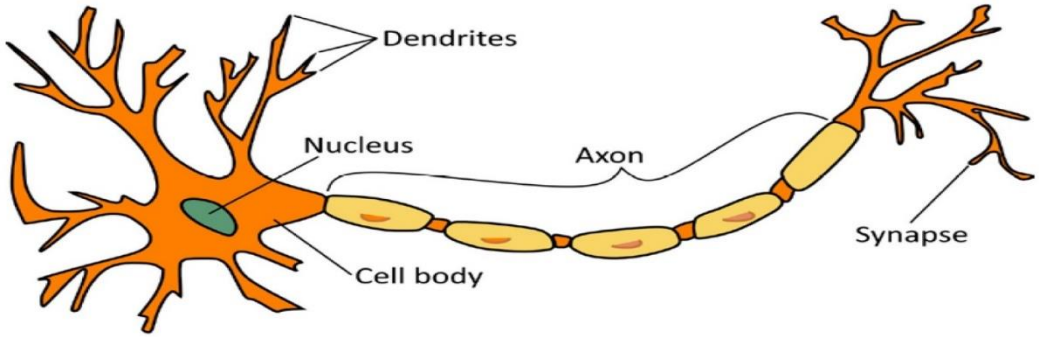


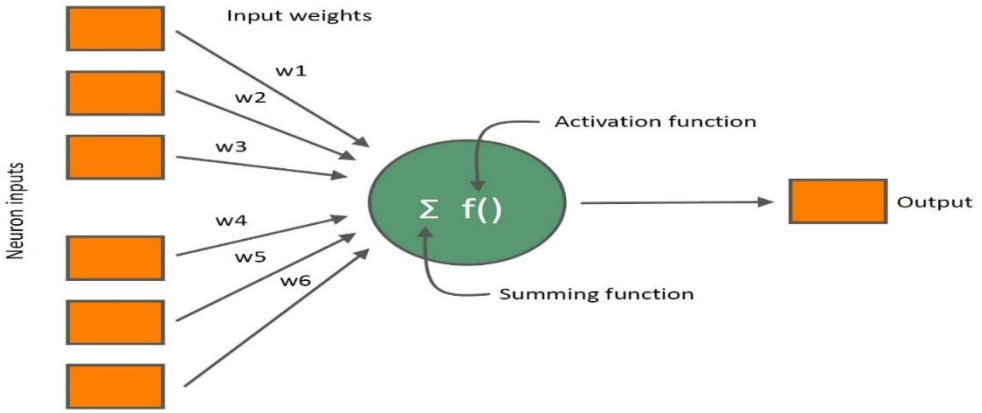
Artificial Neurons

1943

• McCulloch and Pitts simplified mathematical model of neurons and introduced **Neural Networks (NNs)** to model the Brain.



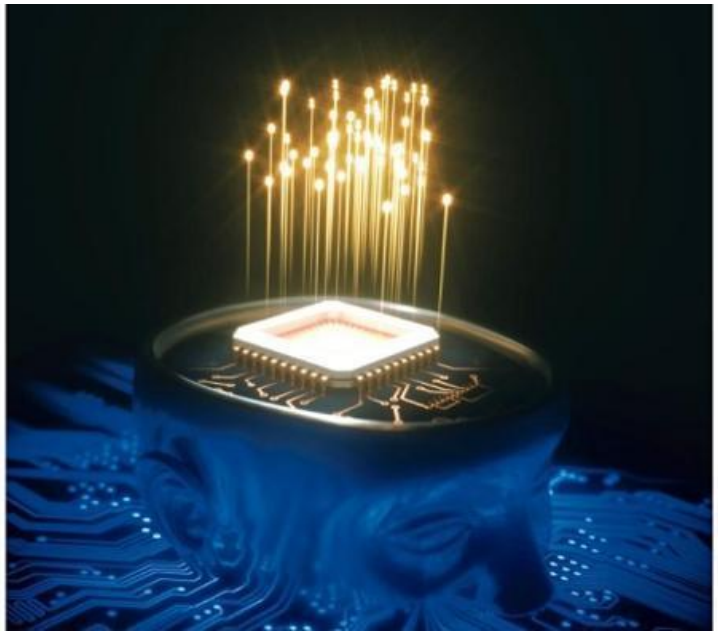




Birth of AI

1956

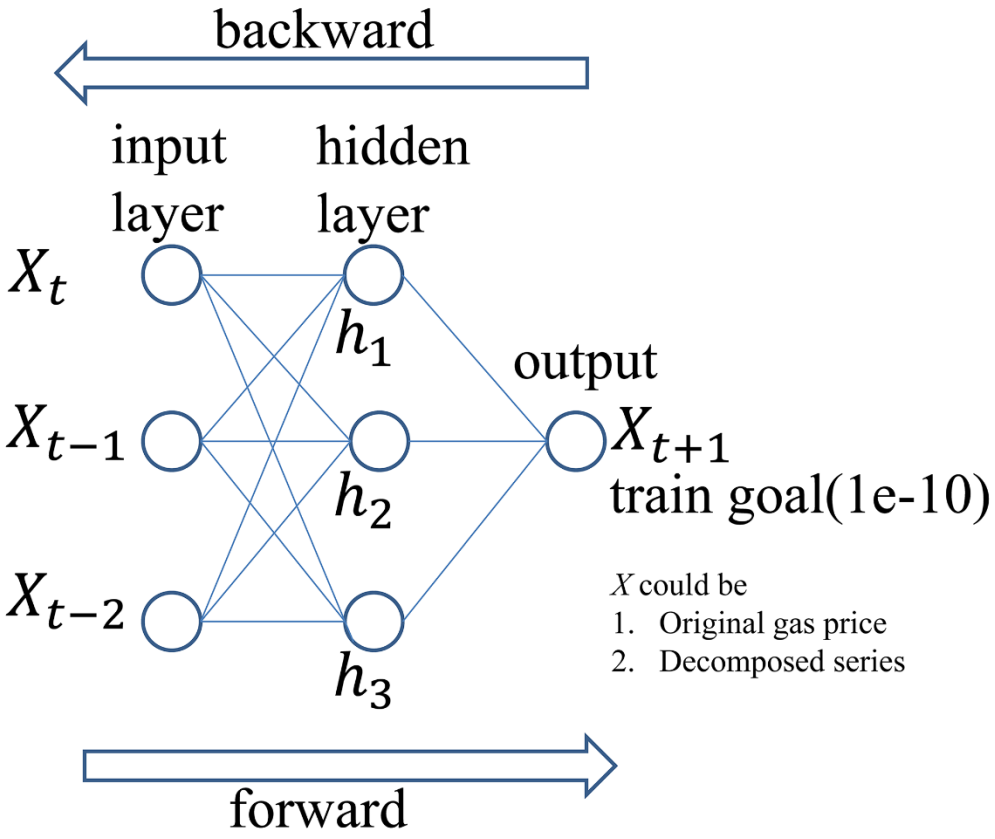
- McCarthy coined the term "**Artificial Intelligence**".
- The term '**Artificial Intelligence**' made its advent at the "Dartmouth Summer Research Project on Artificial Intelligence" conference in Hanover, New Hampshire.



Backpropagation

1960

- Backpropagation Model was developed by Henry J. Kelley.



Backpropagation

- Stuart Dreyfus developed **improved** backpropagation model based on the chain rules.
- However, backpropagation was not used widely till 1985 due to its clumsy and inefficient.

Backpropagation

1965

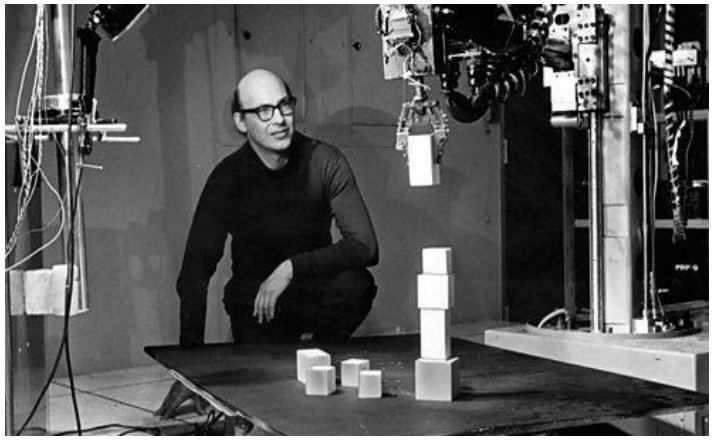
- The Group Method of Data Handling was developed by Alexey Ivankhnenko, as the earliest efforts in developing DL.

Neural Network

1968

- Marvin Minsky and Dean Edmonds built the first **Artificial Neural Network (ANN)**– a computer-based simulation of the way organic brains work.

- The **Stochastic Neural Analog Reinforcement Computer** (SNARC) learned from experience and was used to search a ***maze***, like a rat in a psychology experiment.



Marvin Minsky at MIT in 1968

The First AI ‘Winter’

- In the 1970s, the capabilities of AI programs were limited.
- AI was subject to critiques and financial setbacks.
- AI researchers had **failed** to appreciate the difficulty of the problems they faced.
- Reasons:
  - Limited computer **power**
  - Lack of **data**
  - Commonsense **knowledge** and **reasoning**.
  - Governments and corporations were **losing faith** in AI.

- The **Second AI 'winter’** happened during 1987~1993.

CNN

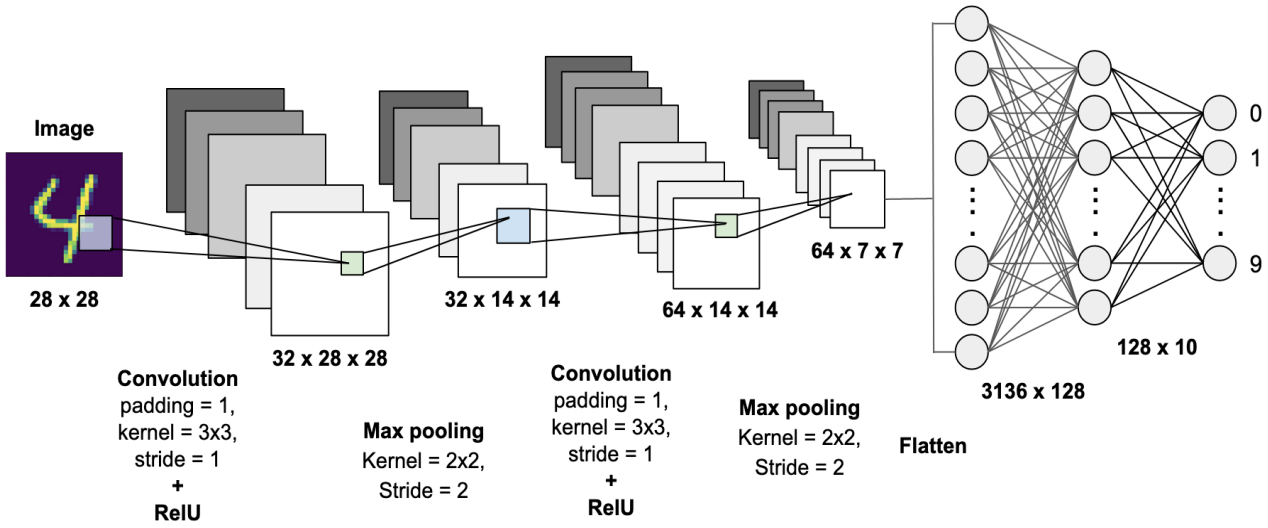
1979

- Kunihiko Fukushima designed ANN with **multiple pooling and convolutional layers**.
- He also developed an ANN, Neocognitron

Handwriting Recognition

1989

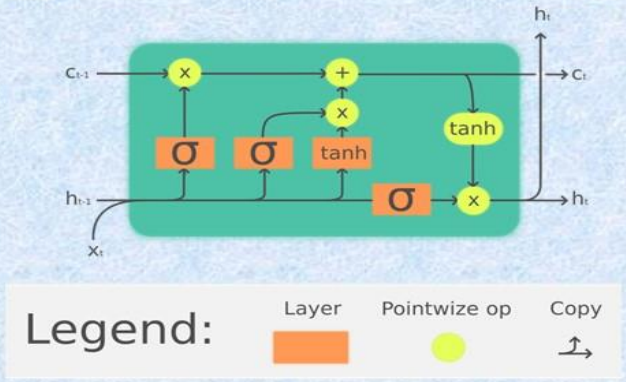
- Yann LeCun invented a machine that can read handwritten digits.



LSTM

1997

- Sepp Hochreiter and Jürgen Schmidhuber publishes a milestone paper on “Long Short-Term Memory” (LSTM).
- LSTM is a type of recurrent neural network architecture which will go on to revolutionize deep learning in decades to come.



Deep Blue

1997

- **Deep Blue** beat world chess champion Garry Kasparov in the first game of a match.
- Kasparov won the 1996 match, but in 1997 an upgraded Deep Blue then won a second match 3½ games to 2½.

DL Birth

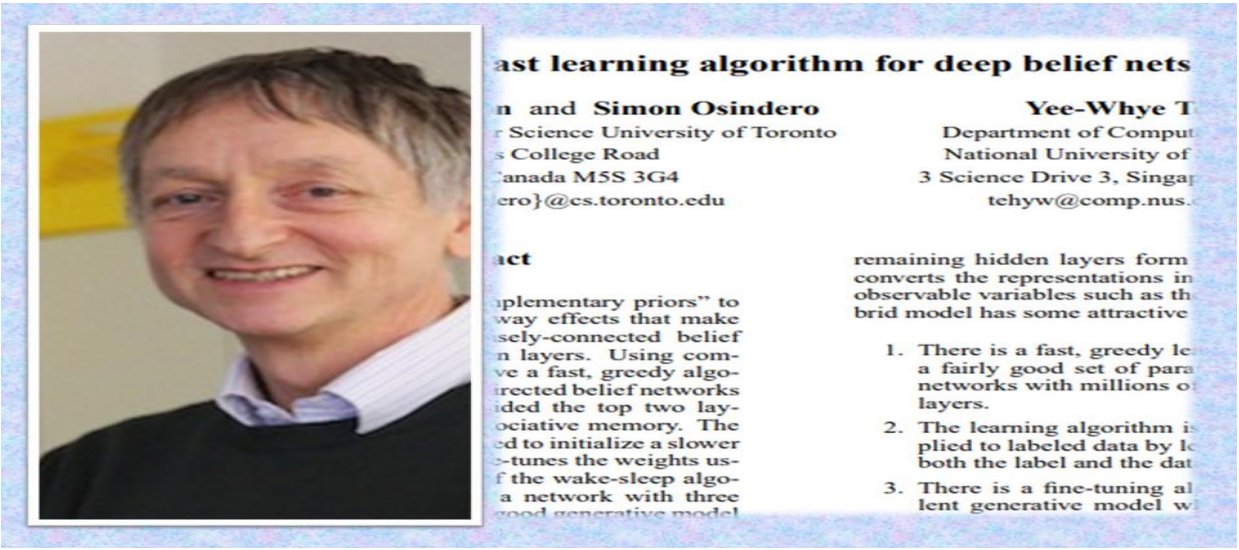
2006

- Hinton coins the term of **Deep Learning**.

Deep Belief Network

2006

- Geoffrey Hinton, Ruslan Salakhutdinov, Osindero and Teh publishes the paper "*A fast learning algorithm for deep belief nets*"
- In which they stacked multiple RBMs together in layers and called them Deep Belief Networks.
- The training process is much more efficient for large amount of data.



ast learning algorithm for deep belief nets

n and Simon Osindero
r Science University of Toronto
s College Road
anada M5S 3G4
ero}@cs.toronto.edu

Yee-Whye T
Department of Comput
National University of
3 Science Drive 3, Singa
tehyw@comp.nus.

act

plementary priors" to
way effects that make
sely-connected belief
n layers. Using com-
ve a fast, greedy algo-
rected belief networks
ded the top two lay-
ociative memory. The
d to initialize a slower
-tunes the weights us-
f the wake-sleep algo-
a network with three
good generative model

remaining hidden layers form
converts the representations in
observable variables such as th
brid model has some attractive

1. There is a fast, greedy le
a fairly good set of para
networks with millions o
layers.
2. The learning algorithm i
plied to labeled data by l
both the label and the dat
3. There is a fine-tuning al
lent generative model w

GPU Revolution

2008

- Andrew NG's group in Stanford starts advocating for the use of GPUs for training DNNs to speed up the training time by many folds.
- This could bring practicality in the field of Deep Learning for training on huge volume of data efficiently.

Watson

2011

- Watson won '**Jeopardy**!'

https://www.youtube.com/watch?v=P18EdAKuC1U

GoogleBrain

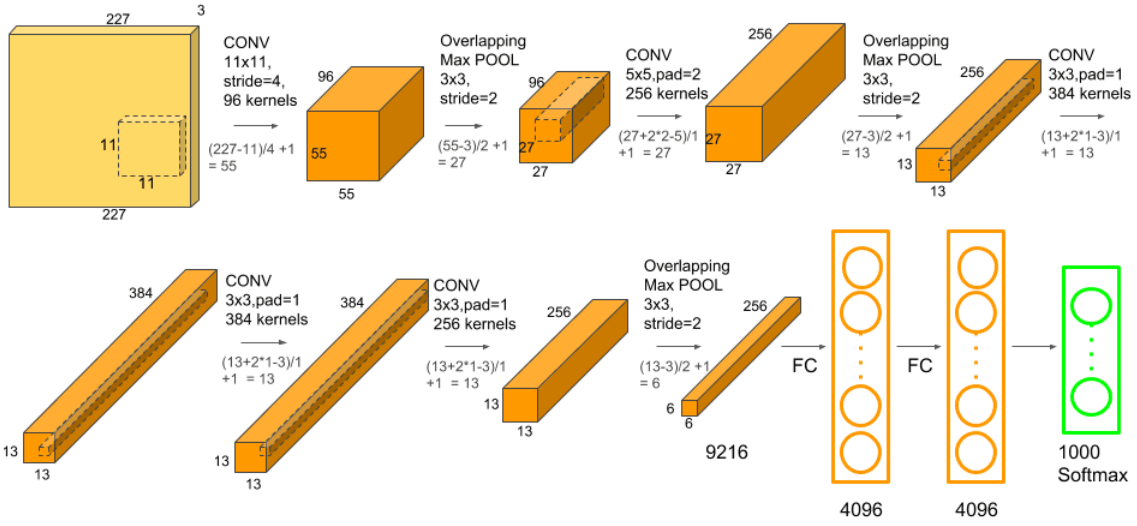
2012

- A DNN created by Jeff Dean, which focused on **pattern detection** in images and videos.

AlexNet

2012

- AlexNet won the ImageNet competition by a large margin in 2012,
- It led to the use of GPUs and **CNNs** in machine learning.
- They also created **ReLU**, which is an activation function that greatly improves efficiency of CNNs.



DeepFace

2014

- A DNN created by Facebook, which they claimed can **recognize** people with the same precision as a human can.

GAN

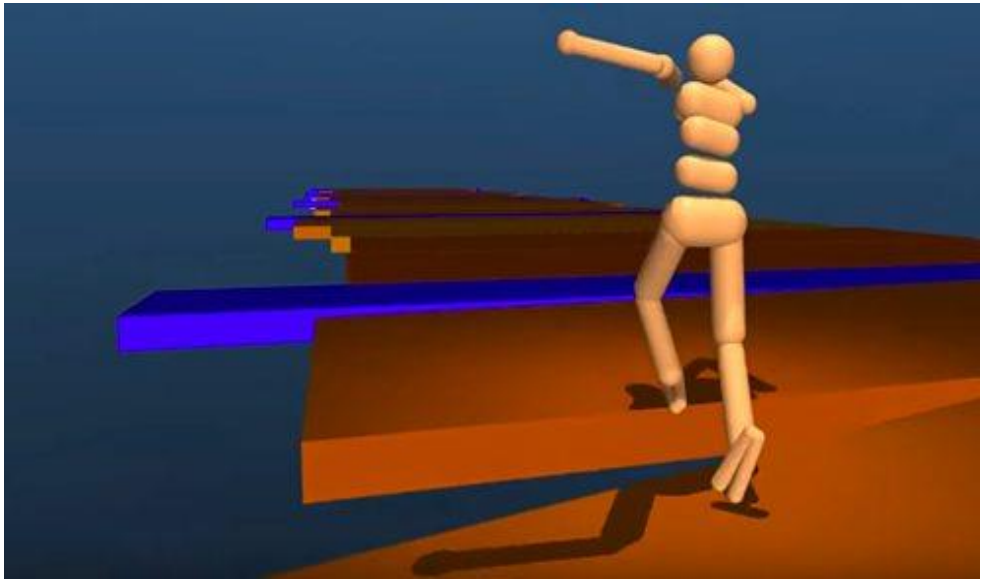
2014

- Ian Goodfellw invites the Generative Adversarial Networks (GANs)

DeepMind

2014

- **Google DeepMind** gained prominence when it developed a neural network that could learn to **play video games**
- By analyzing the behavior of pixels on a screen.



DeepMind artificial intelligence moving an animated figure

AlphaGo beats Lee Sedol

2016

- Developed by **DeepMind** researchers,
- AlphaGo won its first match against a professional in 2015, beat the world's number two player Lee Sedol in March 2016 and the number one player Ke Jie in 2017.
- AlphaGo's neural network is trained by playing both humans and computers, and uses a **Monte Carlo tree search algorithm** to find moves.

Turing Award

2018

- Yoshua Bengio, Geoffrey Hinton, and Yann LeCun wins Turing Award 2018 for their immense contribution in advancements in area of deep learning and artificial intelligence.

Handol

2019

- **HanDol**, developed by South Korea's NHN Entertainment Corp.,
- Took down the 36-year-old master after 180 stones in Lee's final match,
- Held in his hometown of Sinan, 400 kilometers south of Seoul.

State-of-the-art

Computer Vision

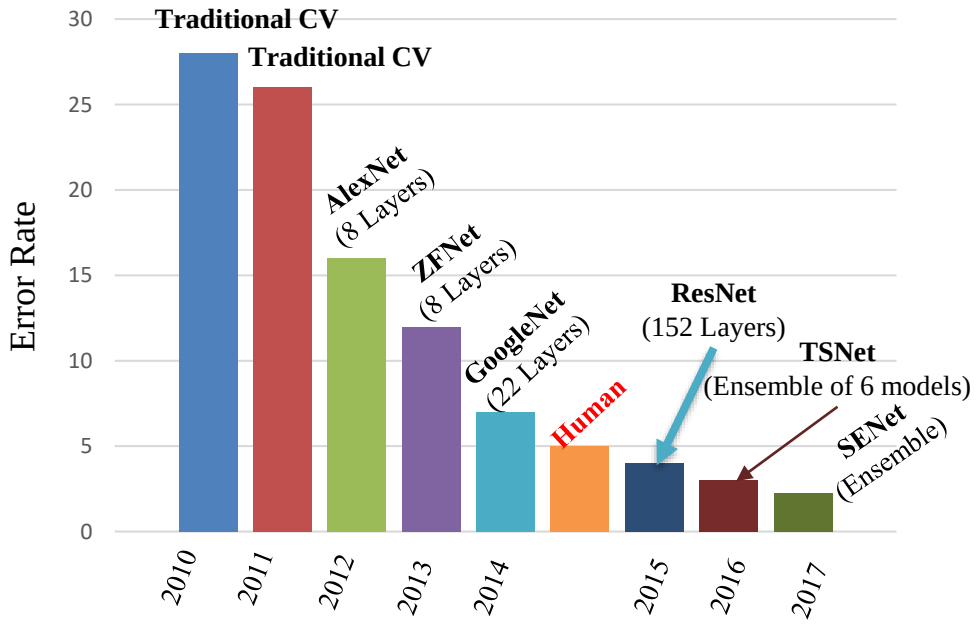
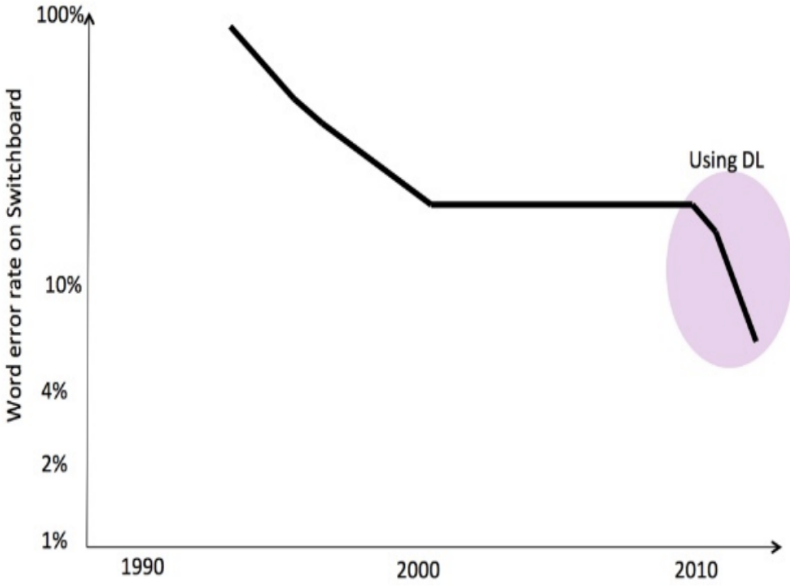


Image classifiers have surpassed human level accuracy.

Speech Recognition

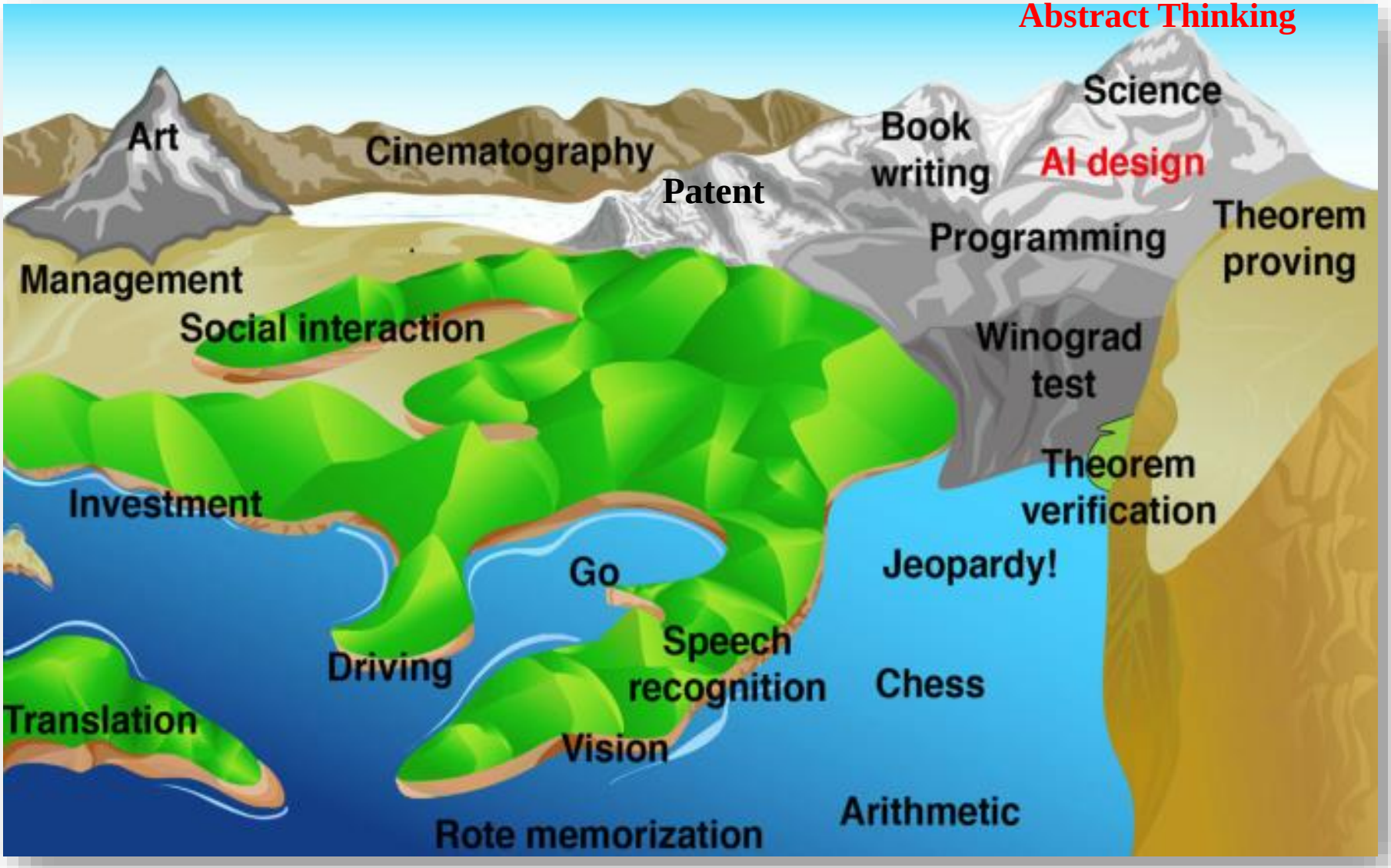


DL improved the accuracy of speech recognition significantly!

State-of-the-art

Game	AI Application	Year of Takeover
Go	Alpha Go	2016
Jeopardy	Watson	2010
Chess	Deep Blue	1997
Checkers	Chinook	1994
Backgammon	BKG	1979

State-of-the-art



Advantages of DL

- Has best-in-class performance on problems that significantly outperforms other solutions in multiple domains, e.g.
  - Speech,
  - Language,
  - Vision,
  - playing games like Go
  - etc.
- Reduces the need for feature engineering, one of the most time-consuming parts of ML practice.
- Can be adapted to new problems relatively easily
- Feature engineering can be automatically executed inside DL model
- Can solve complex problems
- flexible to be adapted to new challenge in the future (or transfer learning can be easily applied)
- High automation.
- Deep learning library (Tensorflow, keras, or MATLAB…) can help users build a deep learning model in seconds (without the need of deep understanding)

Drawbacks of DL

- Requires very large amount of data in order to outperform the other techniques
- Requires high-performance and expensive hardware, e.g. GPUs
- Needs intensive training
- Overfitting if applied into uncomplicated problems
- It’s a blackbox, not straightforward to understand inside each layer
- There is no standard theory to select the best learning tool.
- Determining the topology/flavor/training method/hyperparameters for deep learning is a black art with no theory to guide you.
- What is learned is not easy to comprehend. Other classifiers (e.g. decision trees, logistic regression etc) make it much easier to understand what’s going on.

Assignment

- What is the technique beyond DL?

Office: #432, Dayang AI Center, **Mobile**: 010-8999-8586, **email**: *piran@sejong.ac.kr*