Deep Learning :
A breakthrough in Artificial Intelligence

Muhammad Ghifary, PhD.
Head of Artificial Intelligence

Joelian Samuel, M.Sc.
AI Research Manager

June 2019

What are we going to talk today?

01 Background : Artificial Intelligence and Machine Learning

02 Deep Learning

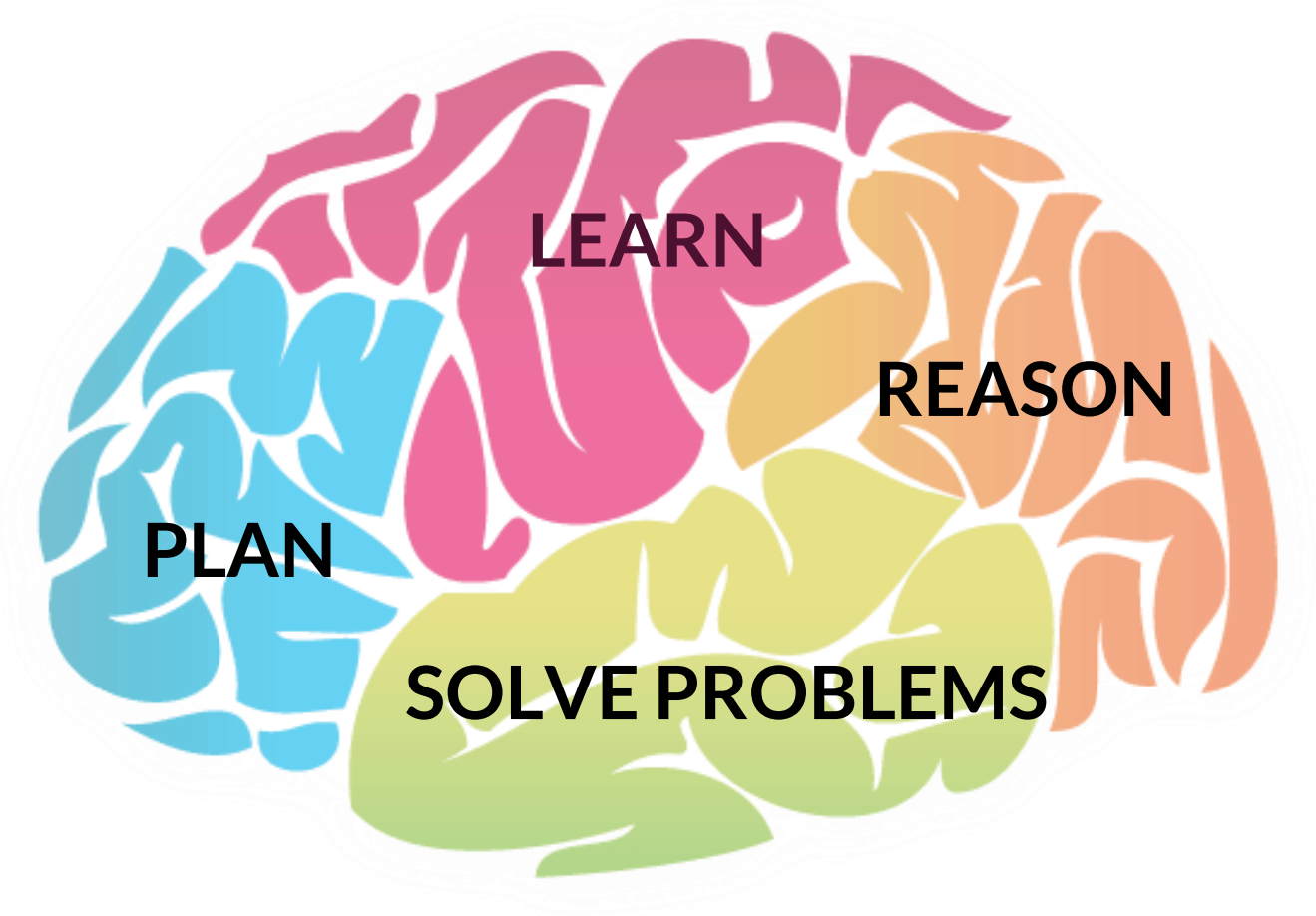
03 Main Principle : Neural Networks

04 Recent Trends & Issues

05 AI in Bukalapak

Background: *AI & ML*

Intelligence

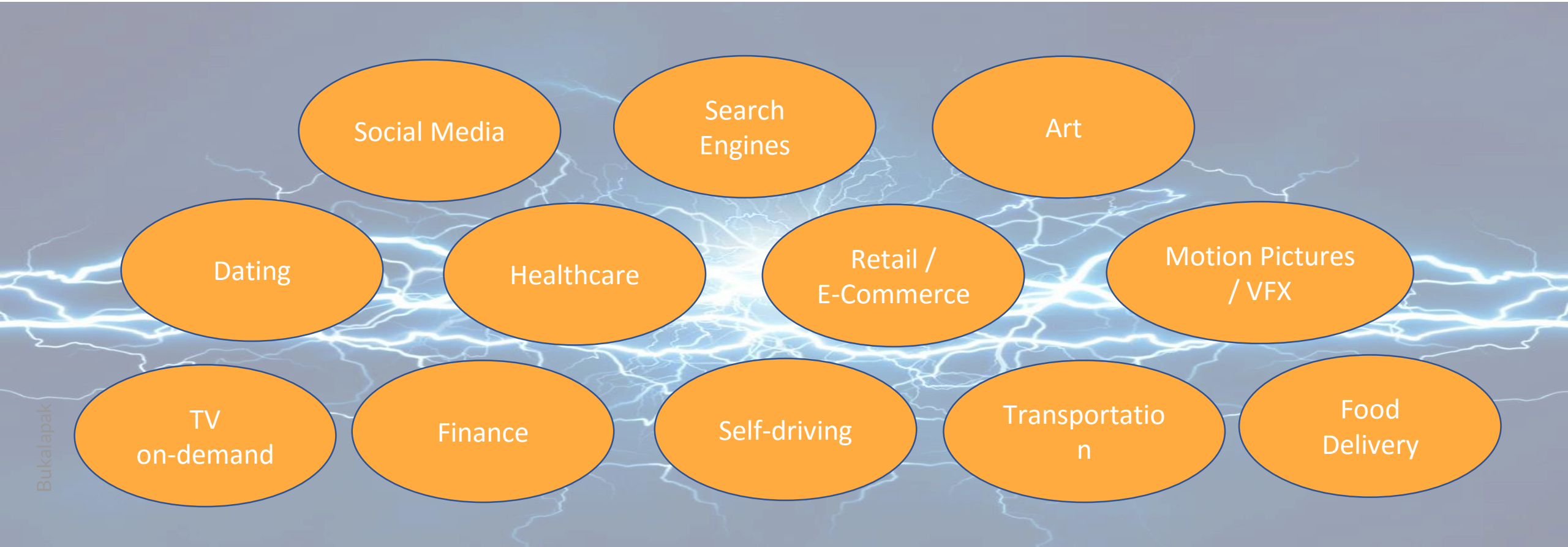


Artificial Intelligence (AI)

The creation of machines / programs that mimic human intelligence

“AI is the new electricity”

- Andrew Ng -



A long time ago, on earth far far away (1956)....

A PROPOSAL FOR THE
DARTMOUTH SUMMER RESEARCH PROJECT
ON ARTIFICIAL INTELLIGENCE

J. McCarthy, Dartmouth College
M. L. Minsky, Harvard University
N. Rochester, I.B.M. Corporation
C.E. Shannon, Bell Telephone Laboratories

August 31, 1955

A Proposal for the

DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

June 17 - Aug. 16

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

The following are some aspects of the artificial intelligence problem:

1) Automatic Computers

If a machine can do a job, then an automatic calculator can be programmed to simulate the machine. The speeds and memory capacities of present computers may be insufficient to simulate many of the higher functions of the human brain, but the major obstacle is not lack of machine capacity, but our inability to write programs taking full advantage of what we have.

2) How Can a Computer be Programmed to Use a Language

It may be speculated that a large part of human thought consists of manipulating words according to rules of reasoning

50 Years later… (2006)

Trenchard More, John McCarthy, Marvin Minsky, Oliver Selfridge, and Ray Solomonoff

Machine Learning

Field of study that gives computers the ability to learn without being explicitly programmed.

Arthur Samuel

A. Samuel (1959). “Some studies in machine learning using the game of checkers”. IBM Journal of Research and Development. 3(3): 210-229.

AI

Artificial Intelligence

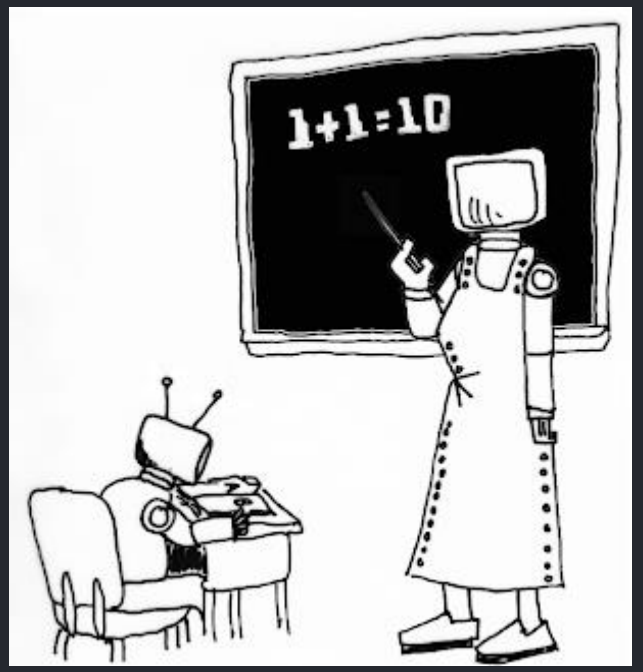
Programming intelligence into computers

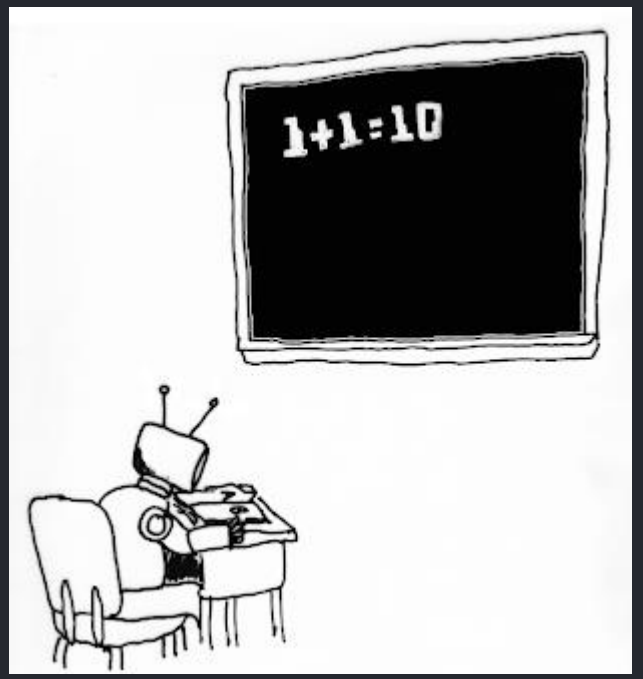
ML

Machine Learning

Programming intelligence into computers,

through learning from data





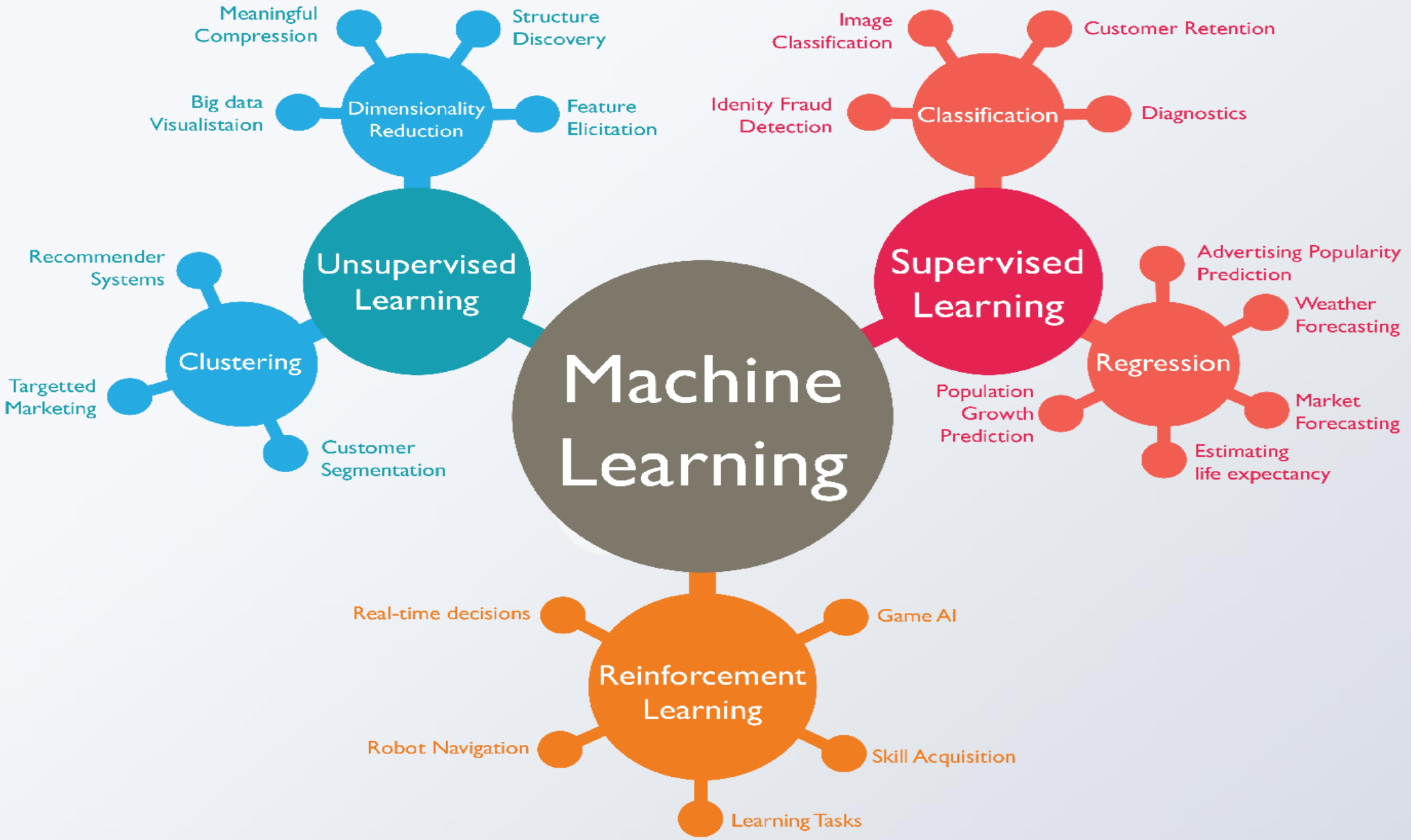
http://prooffreaderswhimsy.blogspot.com/2014/11/machine-learning.html

Machine Learning Paradigms

Supervised Learning

Unsupervised Learning

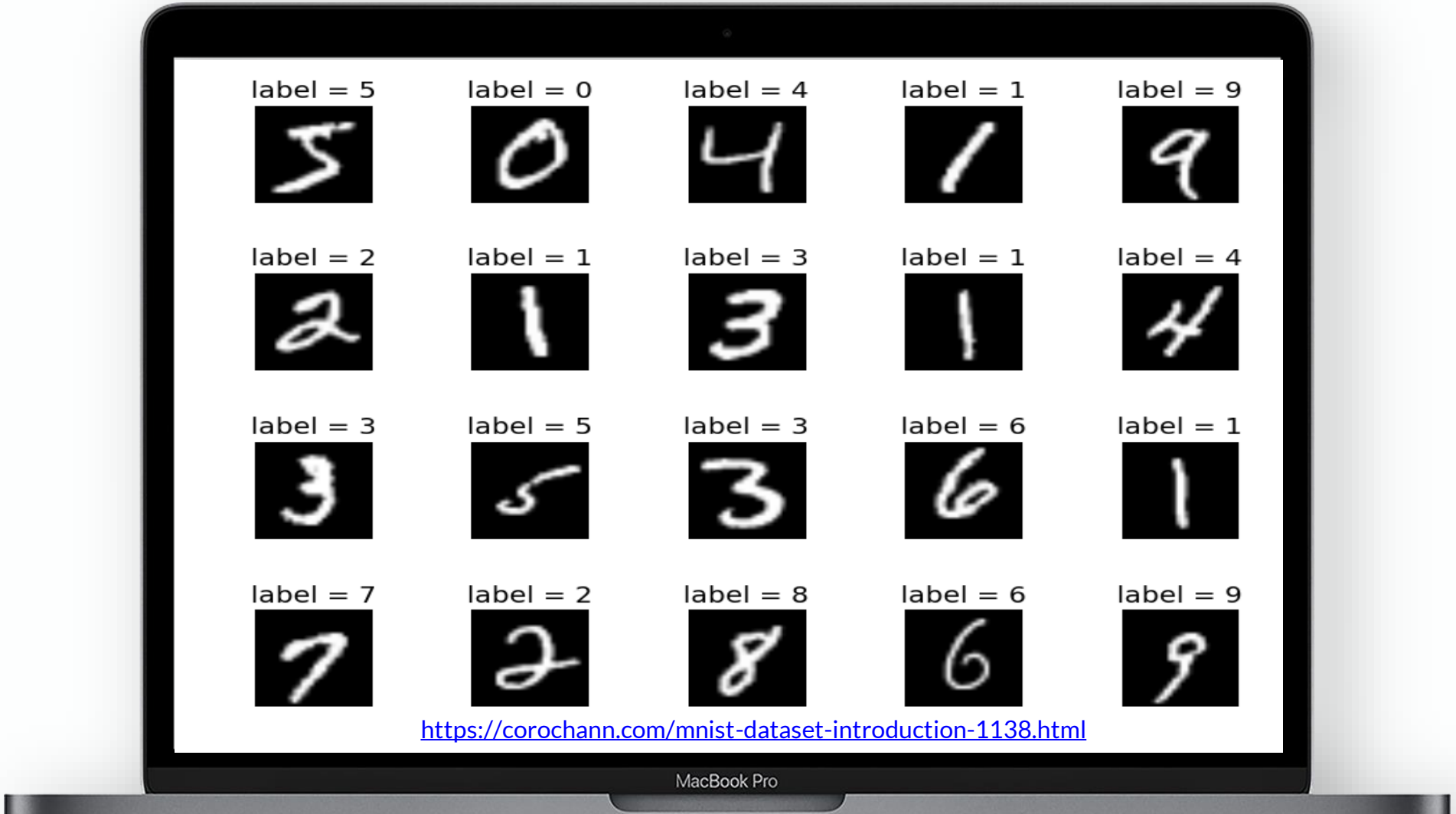
Reinforcement Learning



https://www.linkedin.com/pulse/business-intelligence-its-relationship-big-data-geekstyle/

Supervised Learning

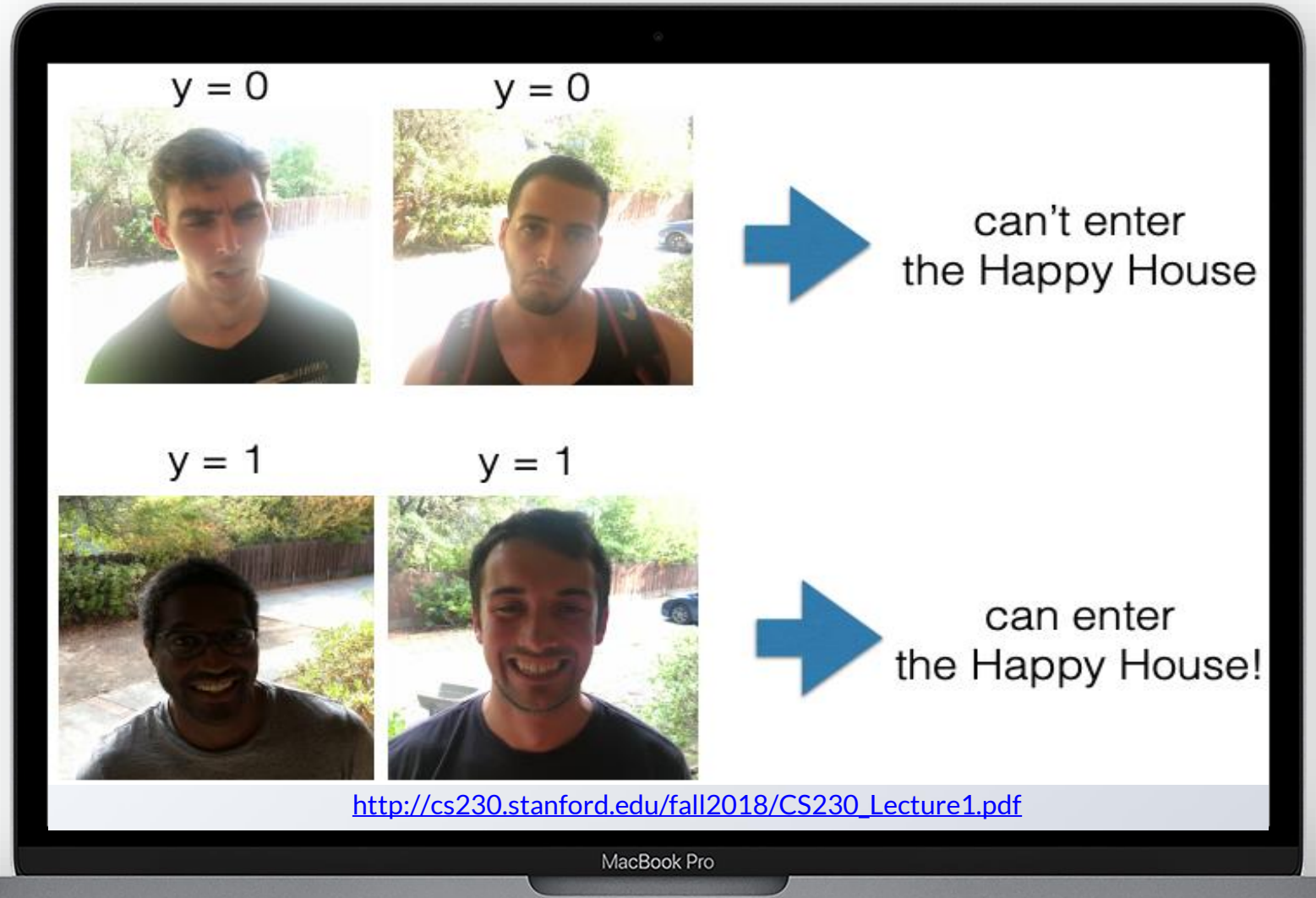
Digit Classification



https://corochann.com/mnist-dataset-introduction-1138.html

Supervised Learning

The Happy House



Supervised Learning

Car Detection

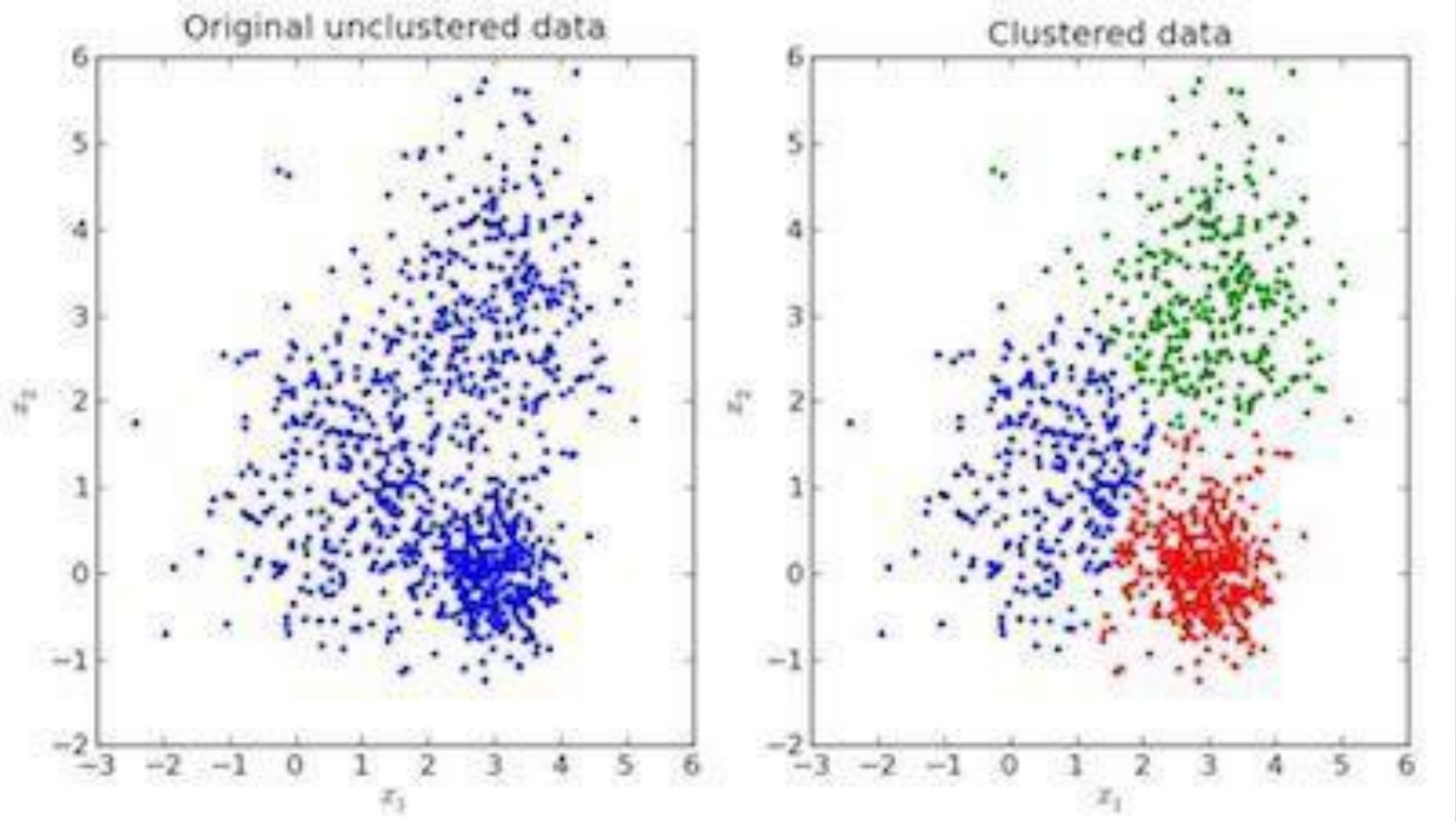
http://cs230.stanford.edu/fall2018/CS230_Lecture1.pdf

and many more ...

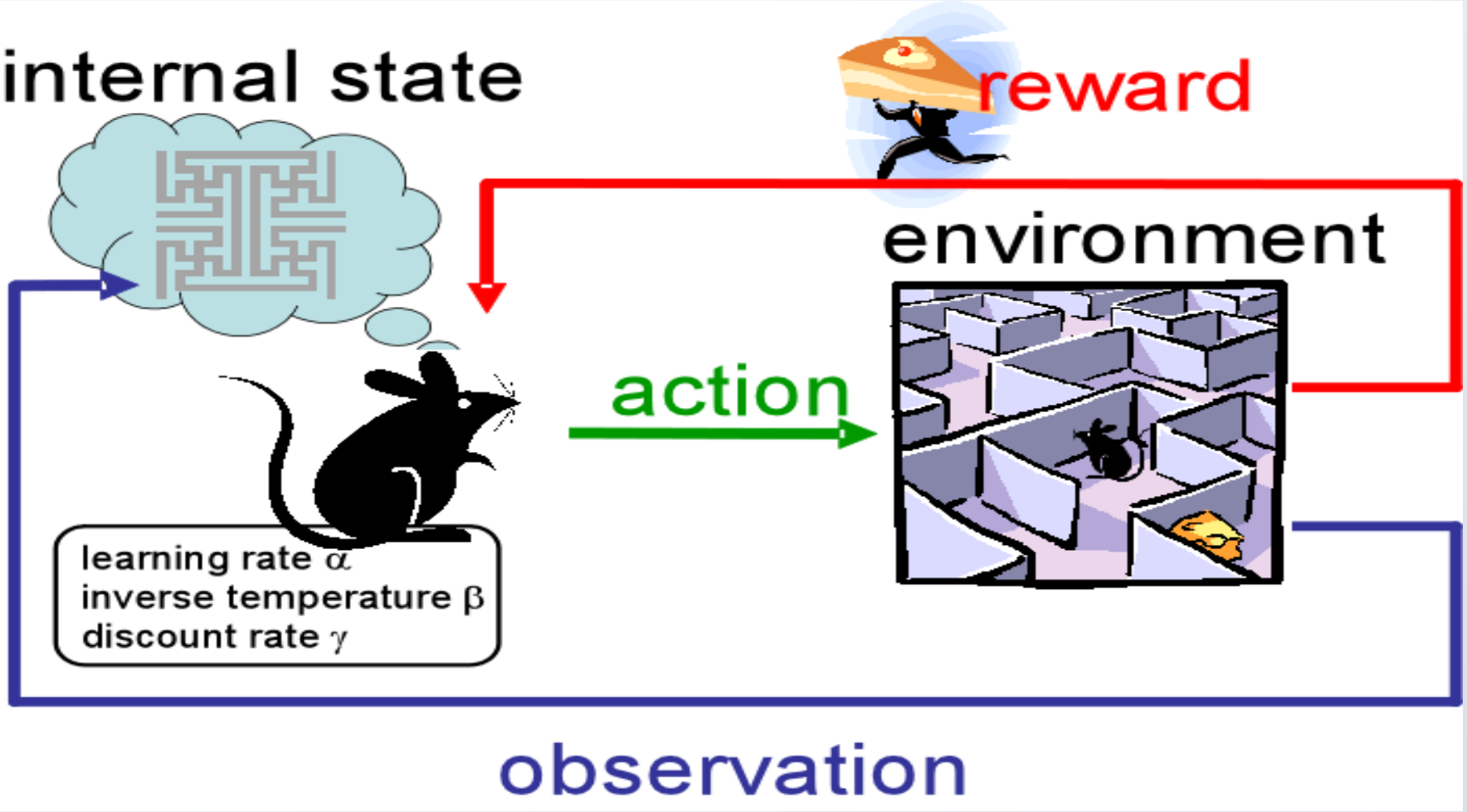
- Cancer detection
- Automatic video surveillance
- Detecting earthquake precursor signals
- Music genre classification
- Visual question answering
- ...

Unsupervised Learning

Clustering



Reinforcement Learning



Tribes in Machine Learning

No.	Klan	Asal usul	Algoritma Master
1.	Symbolists	logika, filsafat	inverse deduction
2.	Evolutionaries	biologi	genetic algorithm
3.	Bayesians	statistik	probabilistic inference
4.	Analogizers	psikologi	kernel machines
5.	Connectionists	neurosains	back-propagation

P. Domingos (2015). “The Master Algorithm”. Basic Books, ISBN 978-0-465-06570-7

What current *AI Can (not) do?*

CS Chatbot

Client: “Saya mau mengembalikan mainan yang tadinya untuk hadiah ulang tahun keponakan saya, karena kirimannya terlambat. Bisa tolong diproses dan sekaligus refund biayanya?”

Template Selection

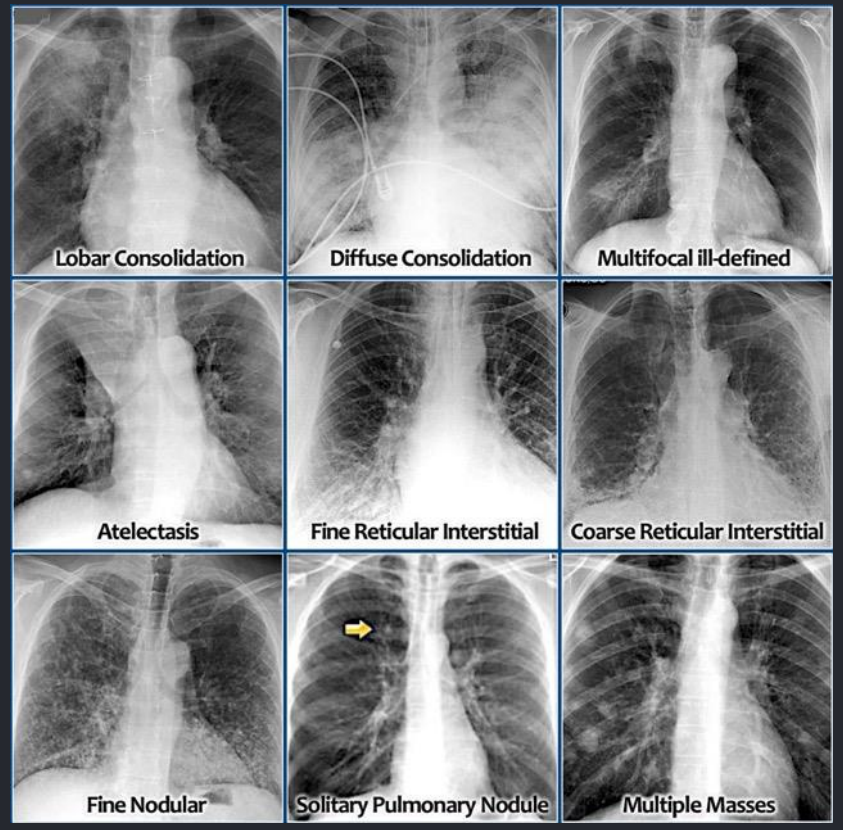
CS: ”Halo kak. Saya dengan Bella. Terima kasih atas pesannya. Saya arahkan permintaannya ke pengajuan refund"

Text Generation

CS: ”Halo kak. Saya dengan Bella. Mohon maaf atas keterlambatan pengirimannya. Semoga ulang tahun keponakan berjalan lancar dan meriah! Baik kak, saya bisa bantu untuk ...."

Self-driving

Deep Learning



Diagnose pneumonia from ~10,000 labeled images

Diagnose pneumonia from 10 images of a medical textbook chapter explaining pneumonia

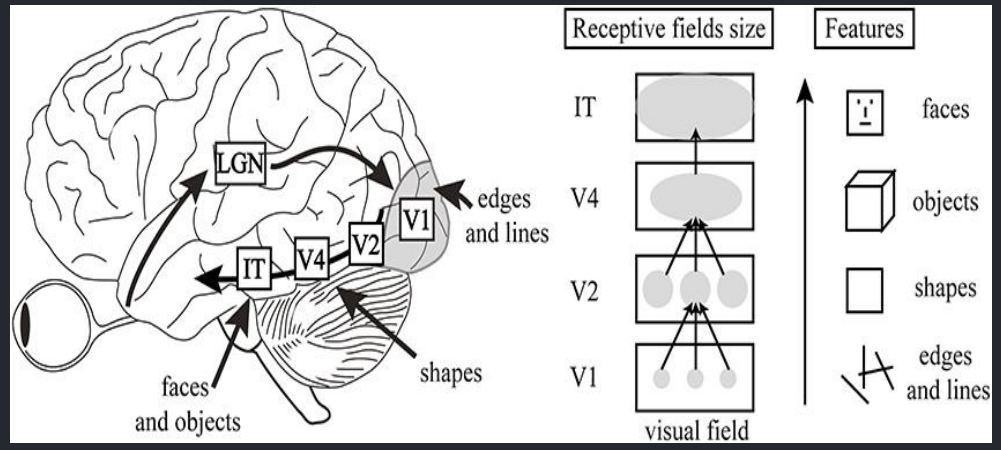
Bukalapak

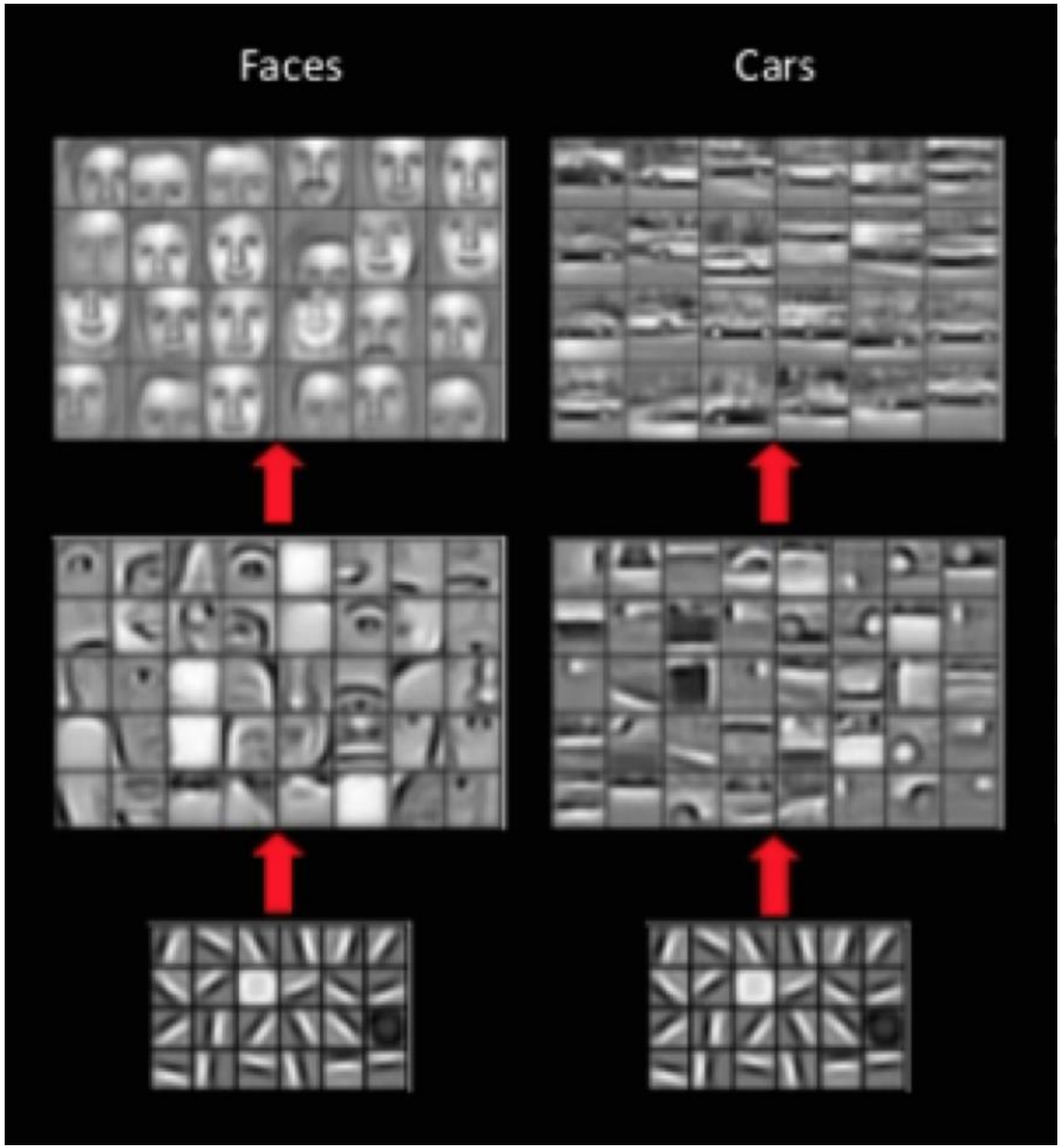
IN CS, IT CAN BE HARD TO EXPLAIN THE DIFFERENCE BETWEEN THE EASY AND THE VIRTUALLY IMPOSSIBLE.

Deep Learning

What's Deep Learning?

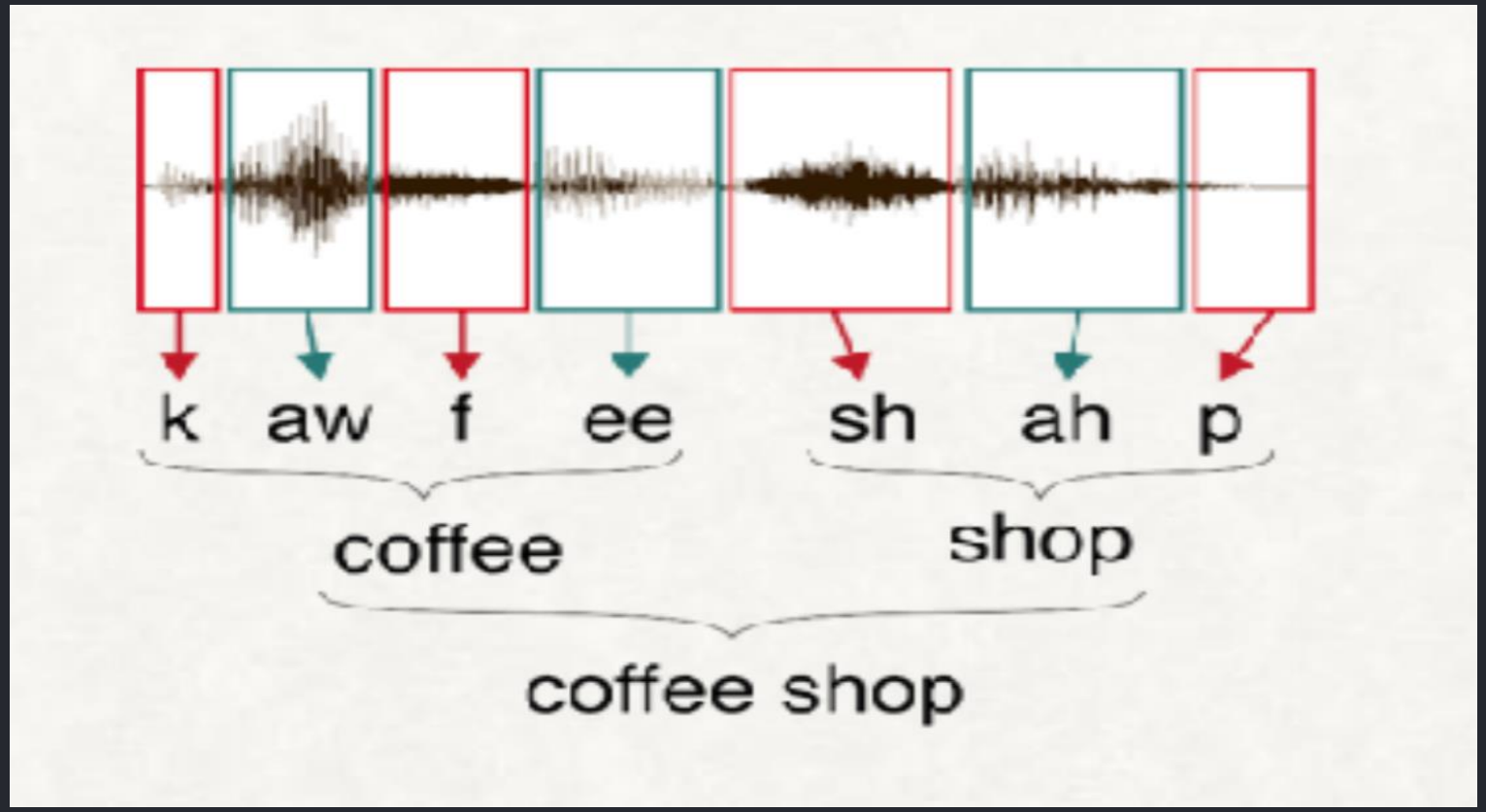
(Hierarchical) representation / feature learning

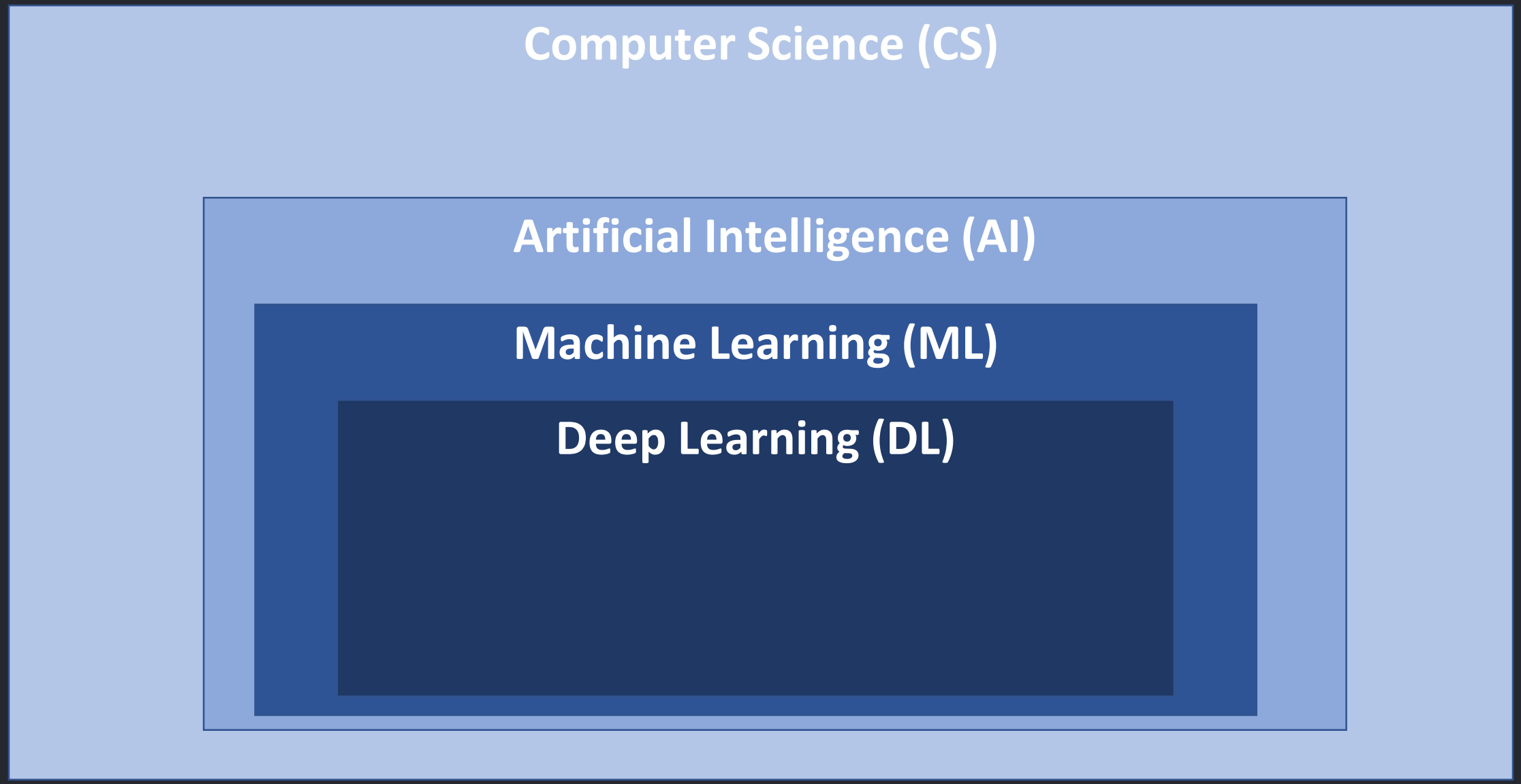




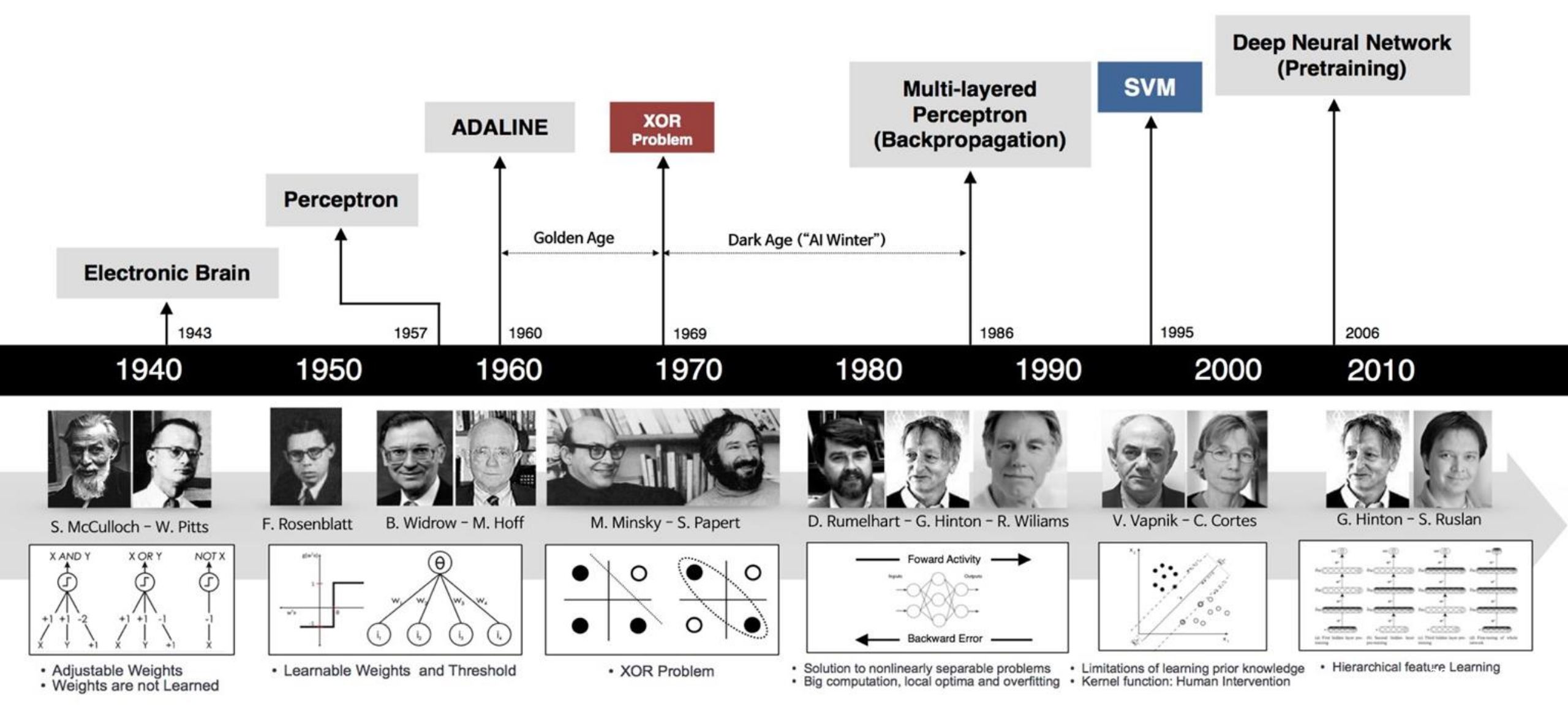
http://web.eecs.umich.edu/~honglak/icml09-ConvolutionalDeepBeliefNetworks.pdf

What's Deep Learning?




Computer Science (CS)
Artificial Intelligence (AI)
Machine Learning (ML)
Deep Learning (DL)

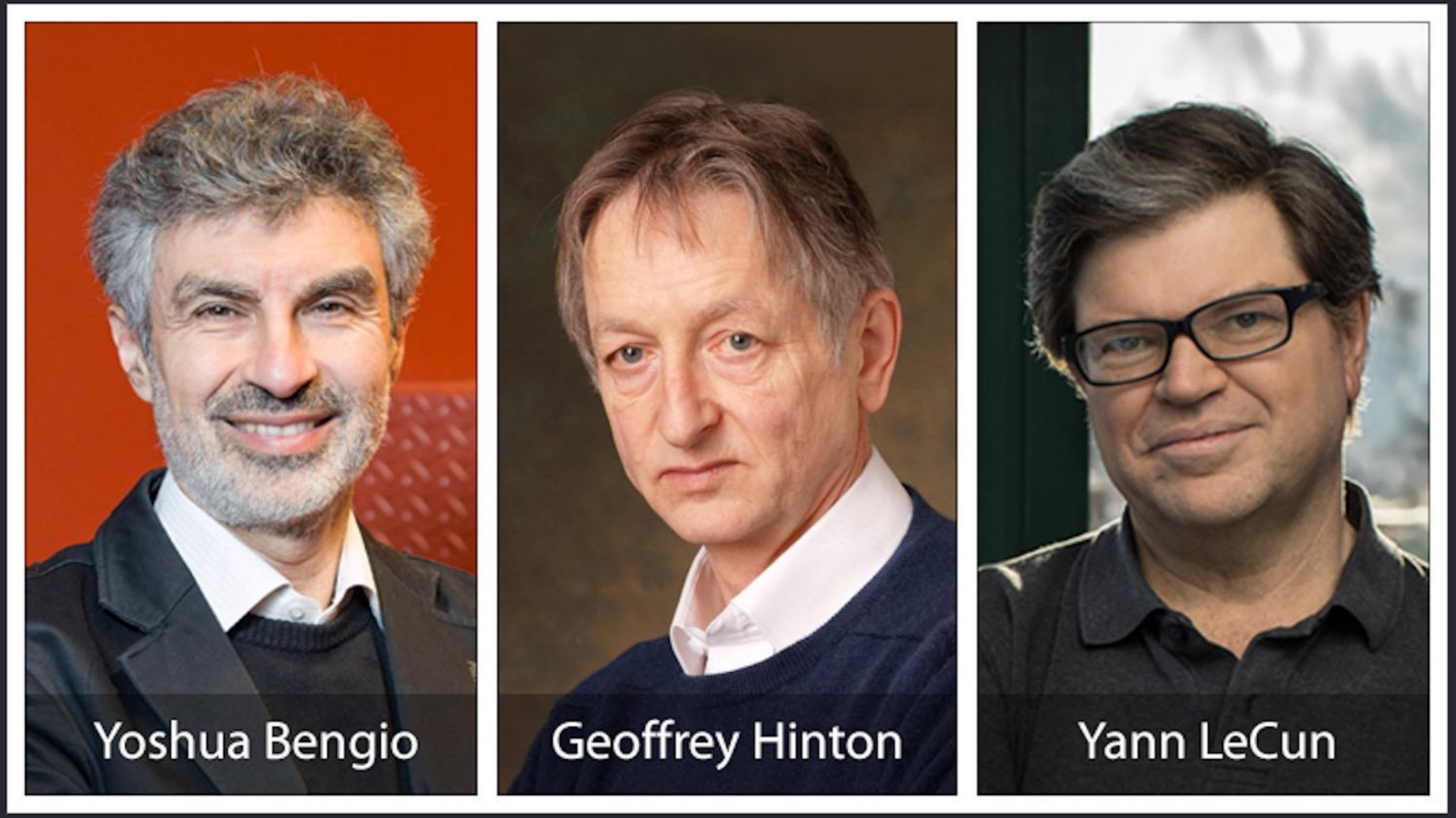
A Brief History of Deep Learning



A Brief History of Deep Learning (2)

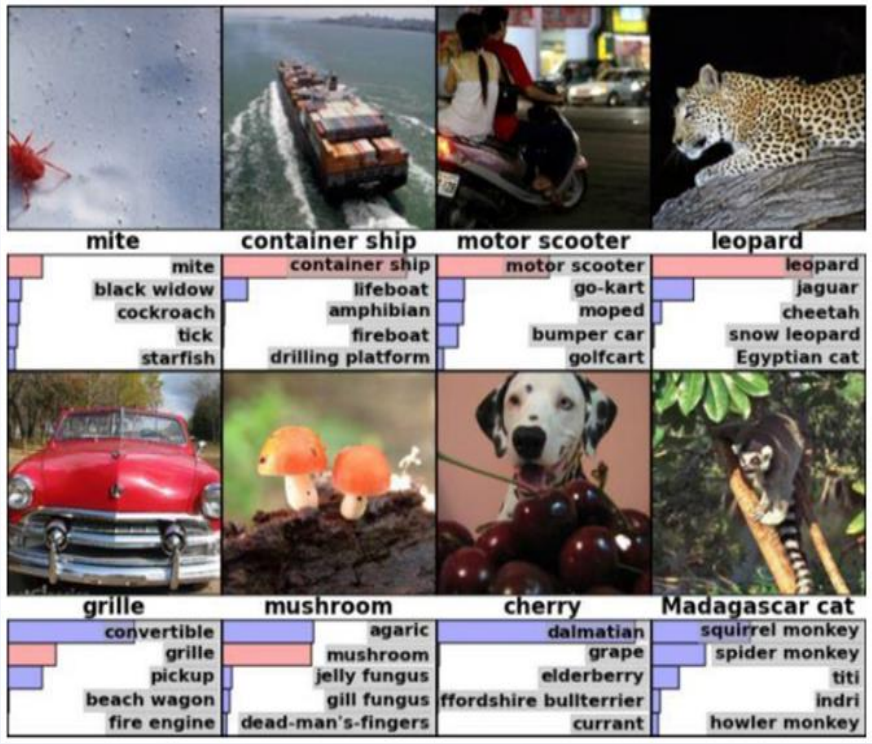
- 2012: AlexNet (object classification)
- 2013: Word2Vec (NLP), R-CNN (object detection & localization)
- 2014: Generative Adversarial Networks, Variational Auto-encoders (generative model)
- 2015: Residual Networks (ResNet) (object classification)
- 2016: AlphaGo (game)
- 2017: AlphaGo Zero (game), Transformer (language translation)
- 2018: BERT (NLP)
- 2019: ?

2018 ACM A. M. Turing Award Laurates



ImageNet 2012

A breakthrough in computer vision



2012

Team	Top-5 Error (%)
AlexNet	**15.315 %**
ISI	26.172 %
OXFORD_VGG	26.979 %
XRCE/INRIA	27.058%
University of Amsterdam	29.576%
LEAR-XRCE (non-ConvNet)	34.464%

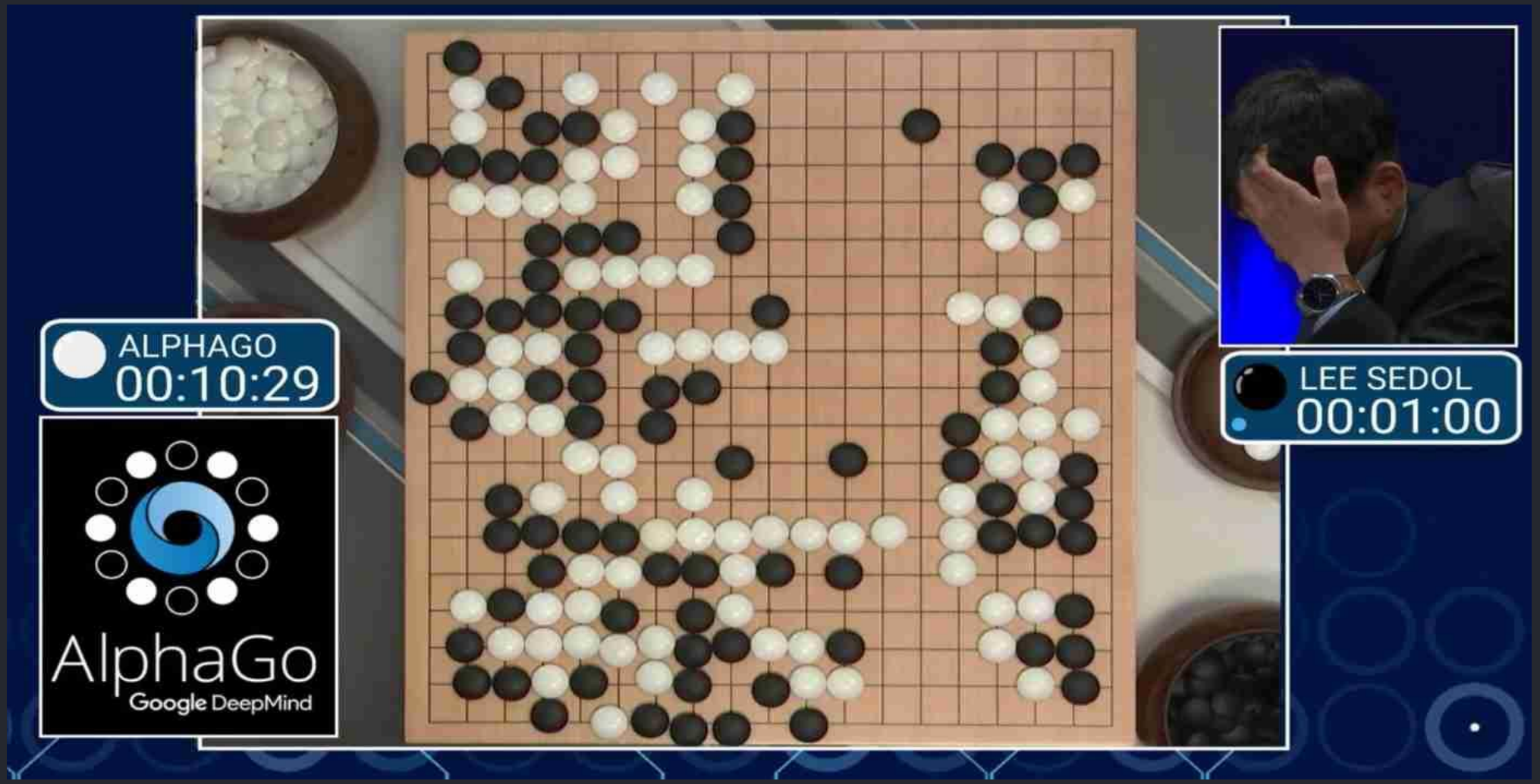
Krizhevsky et al. 2012. “ImageNet Classification with Deep Convolutional Neural Networks”. NIPS 2012

ImageNet 2014 - 2015

Better than human

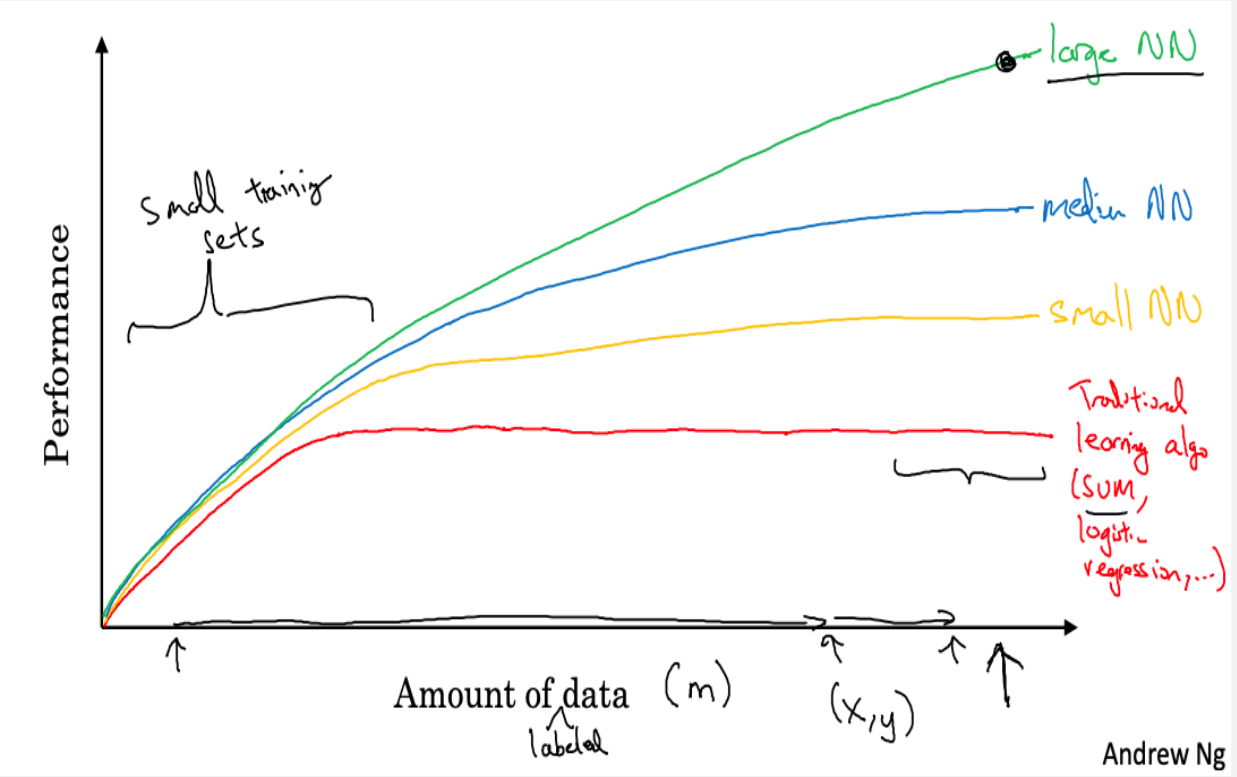
Year	Method	Top-5 Error (%)
2014	VGG	7.32 %
2014	GoogLeNet	6.67 %
	Human	**5.1 %**
2015	**MSRA, PReLU**	**4.94 %**

AlphaGo (2016 - 2017)

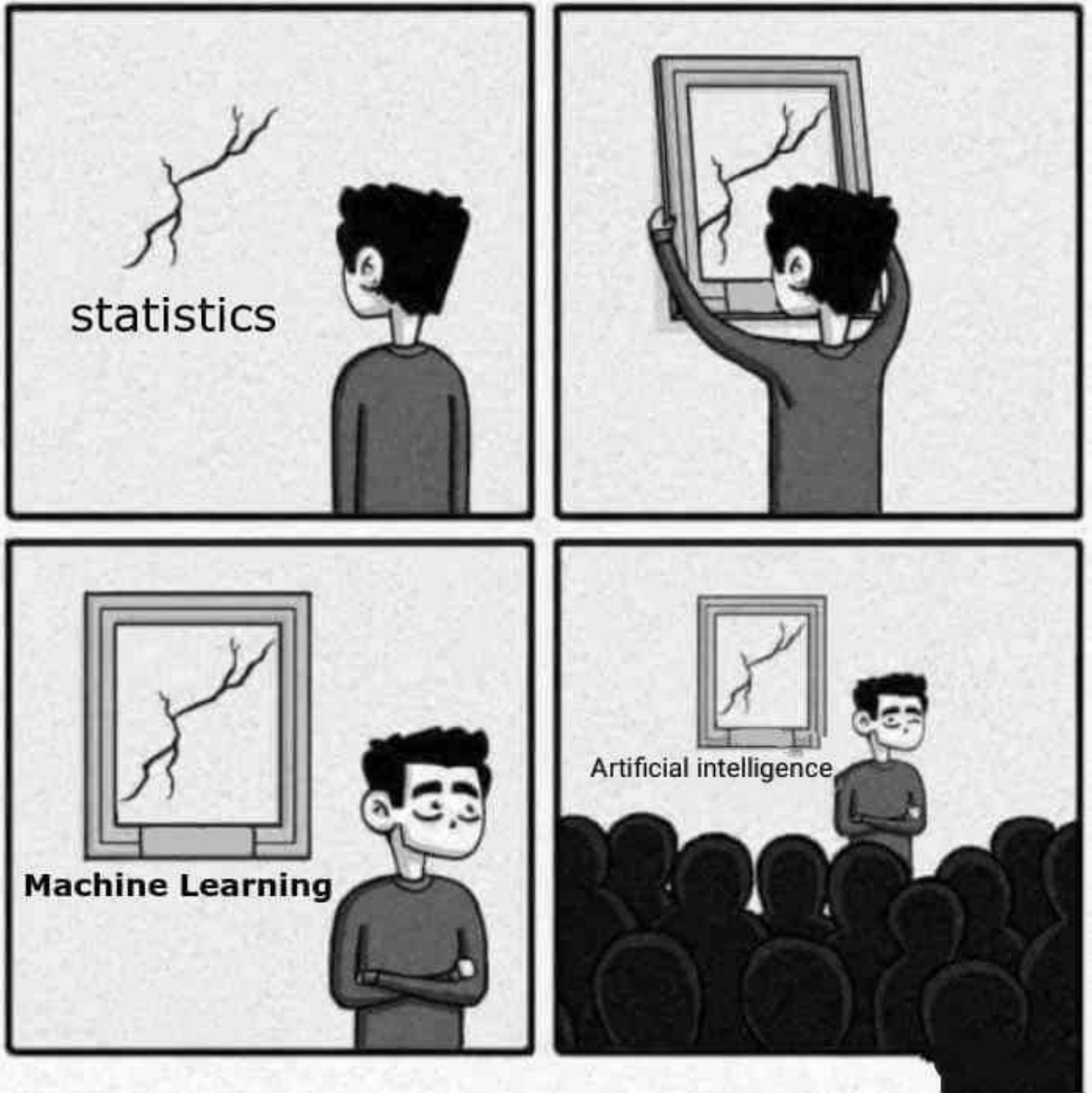


Why Deep Learning?

1. Big Data
2. Scalability: more data, better
3. Hardware Advancement



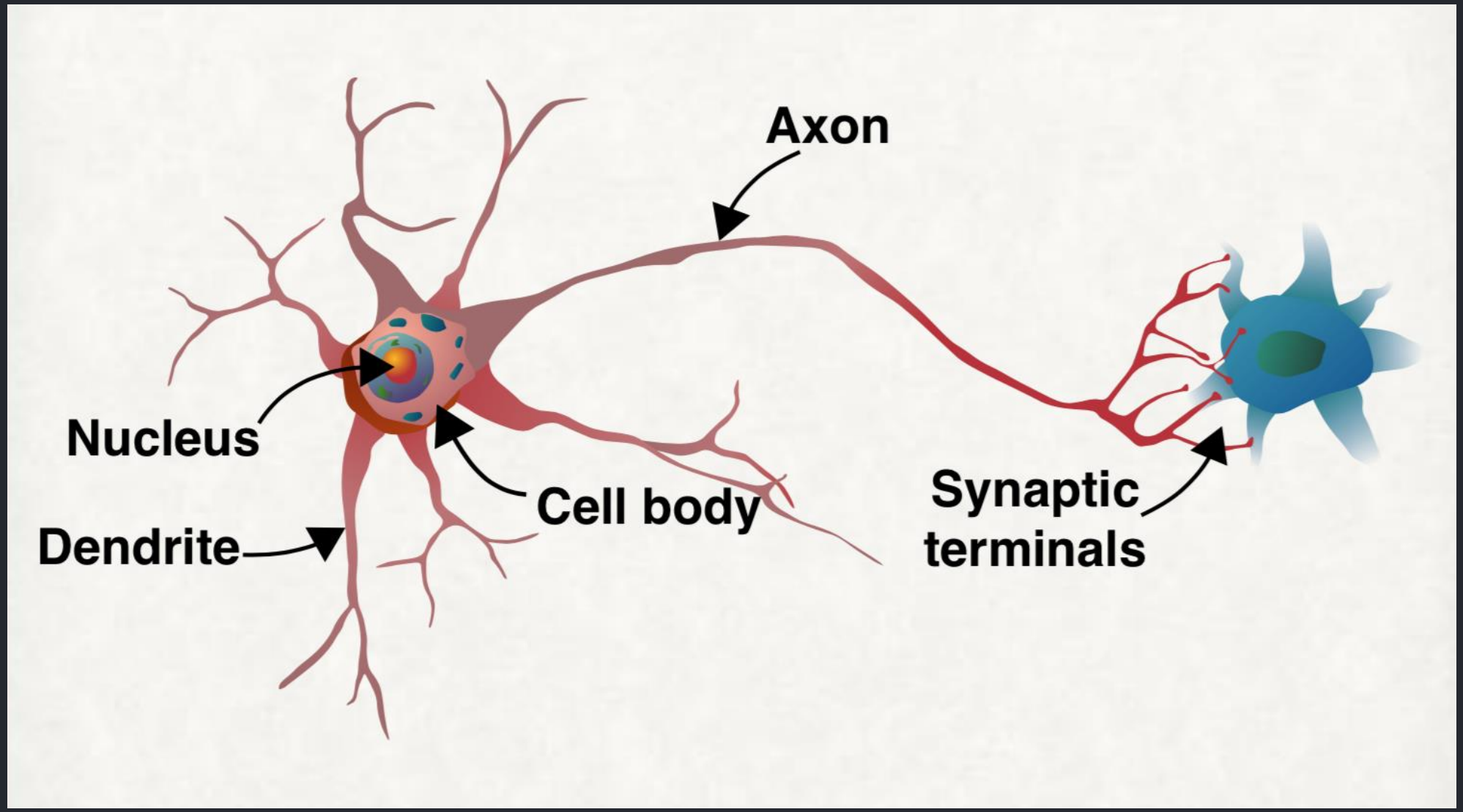
http://cs230.stanford.edu/files/C1M1.pdf


statistics
Machine Learning
Artificial intelligence

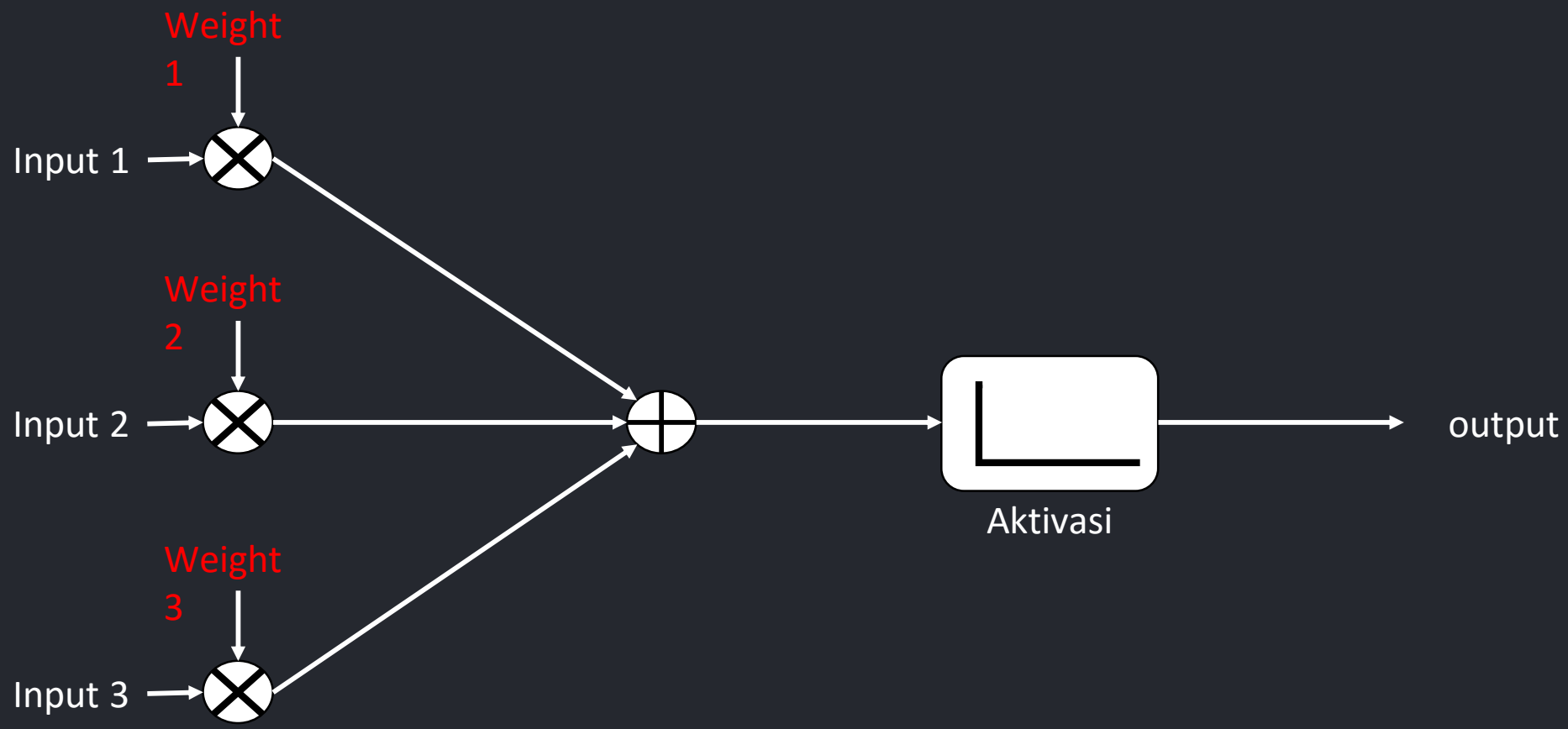
Main Principles:

Neural Networks

Neurons

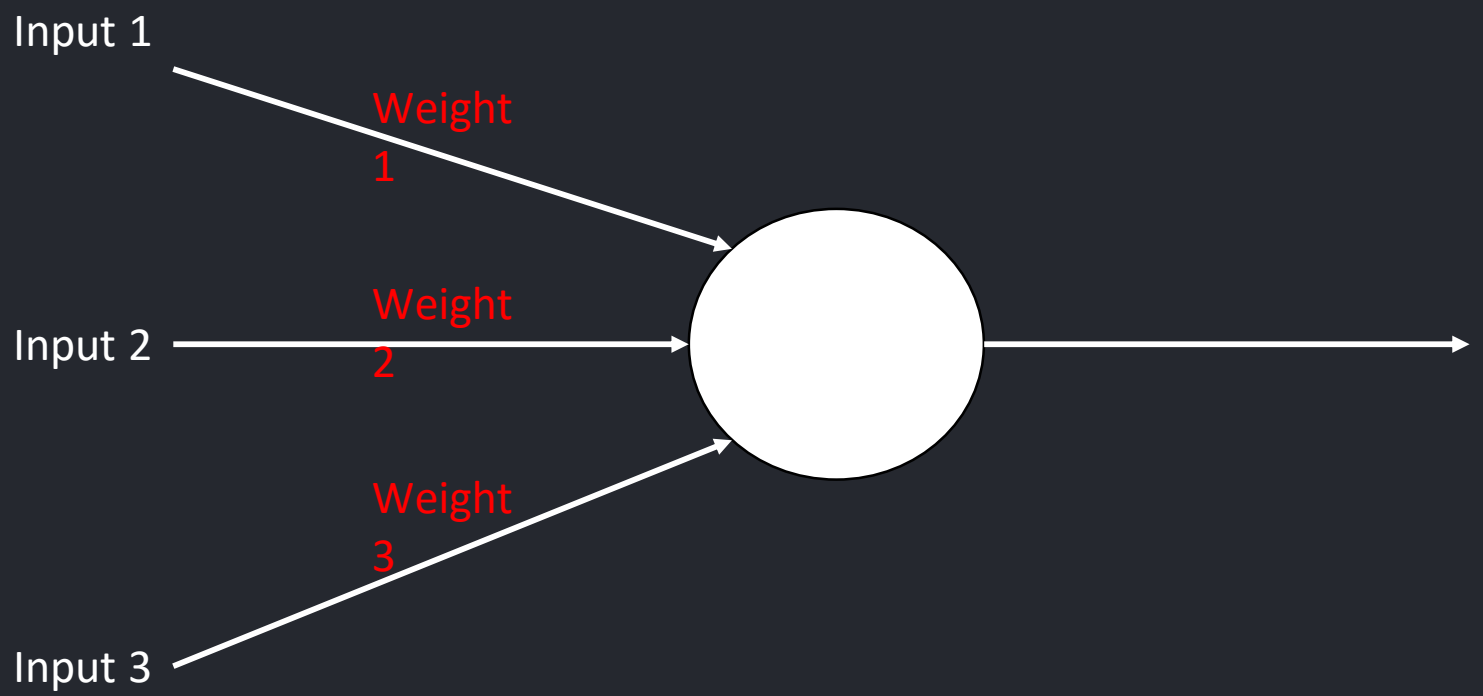


Perceptrons



output = Aktivasi (input_1 x weight_1 + input_2 x weight_2 + input_3 x weight_3)

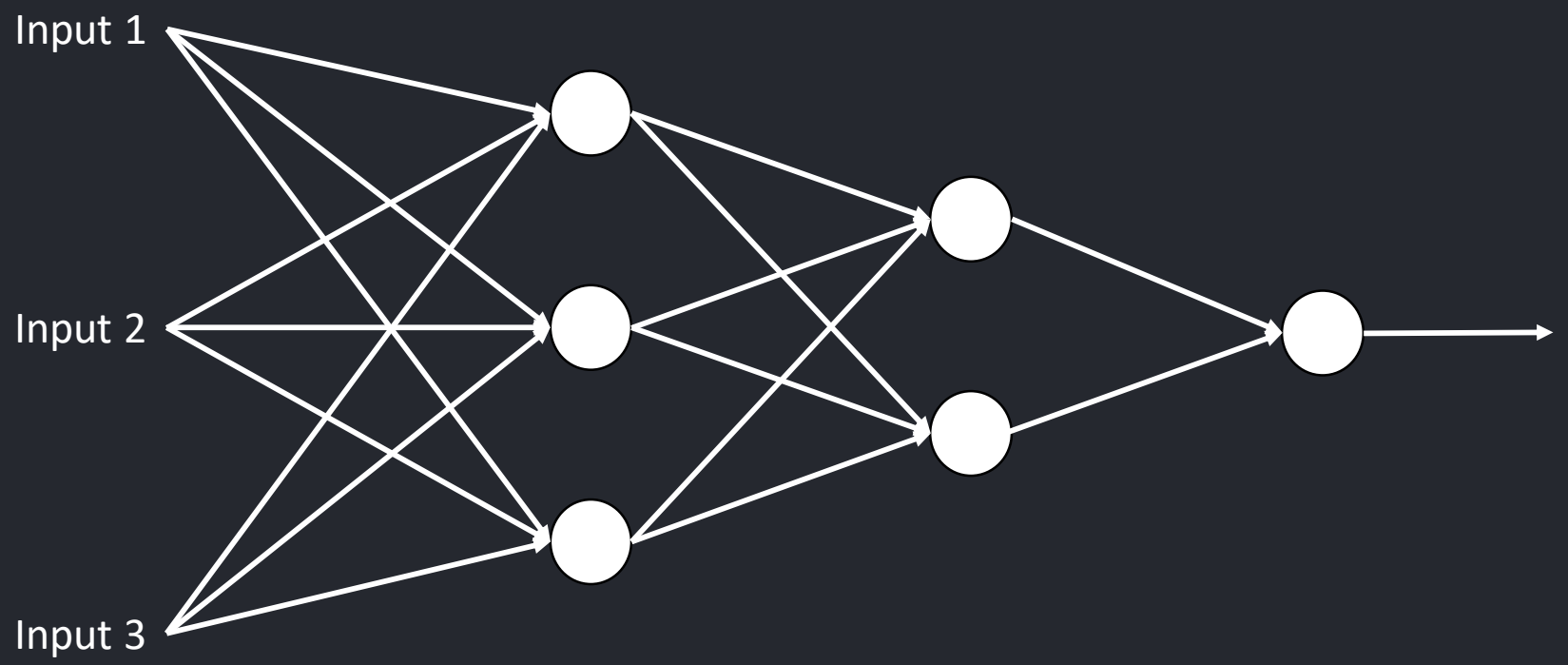
Perceptrons



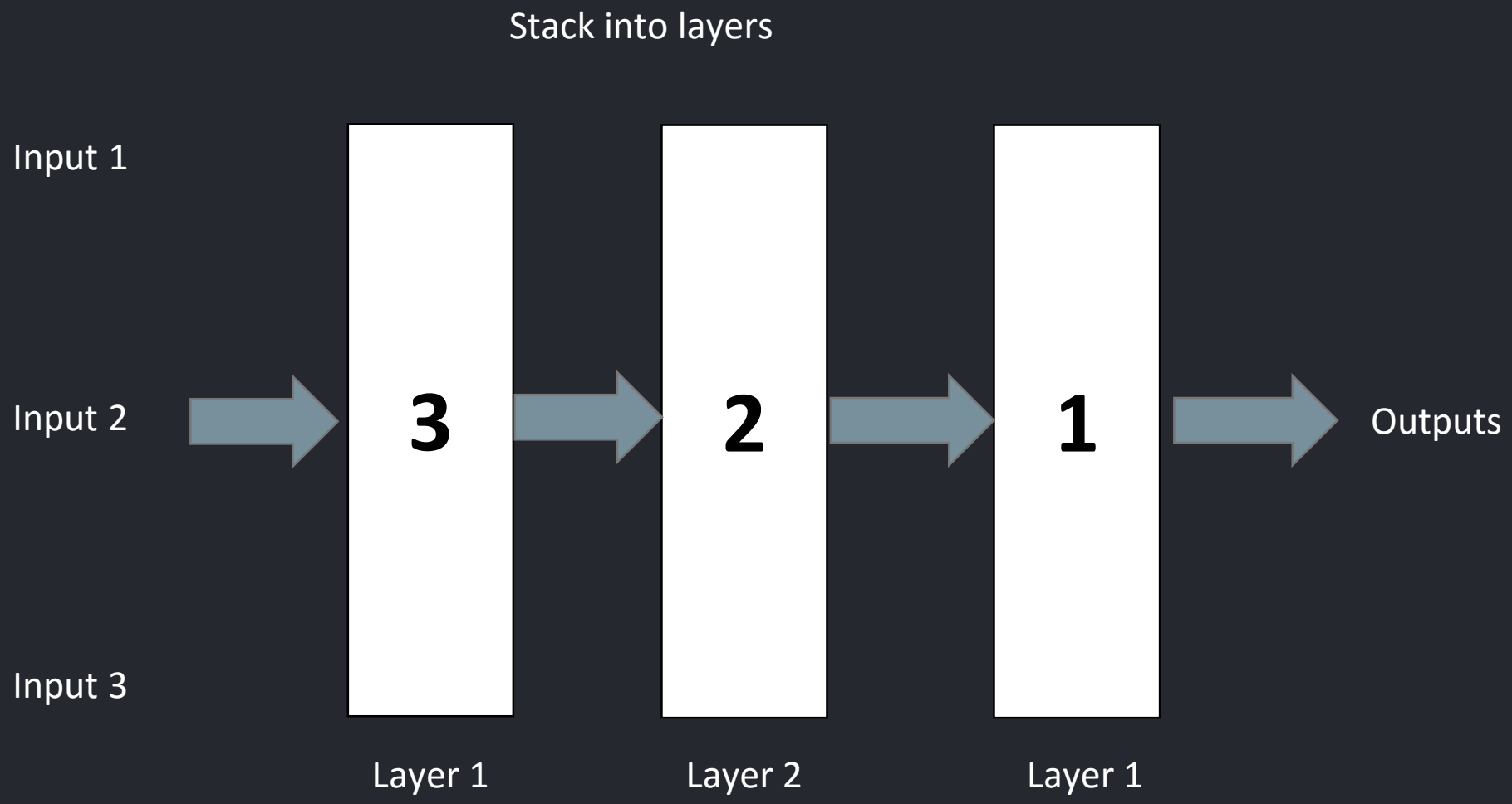
output = Aktivasi (input_1 x weight_1 + input_2 x weight_2 + input_3 x weight_3)

Multi-layer Perceptrons: Neural Networks

Stack into layers



Multi-layer Perceptrons: Neural Networks



How do neural networks learn?

Learning from mistakes by changing the weights

http://leecaraher.com/mistakes-important-learn-success/

Supervised Learning Template

Two Phases

Training
(Offline)

Prediction

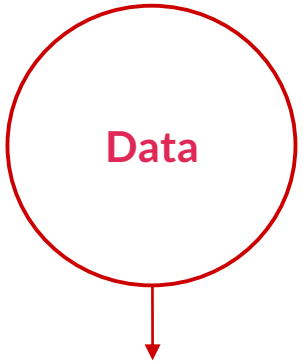
Supervised Learning Components

01 Model / Function

02 Evaluation / Objective

03 Training / Optimization

Supervised Learning Components



01 Model / Function

Function to be trained

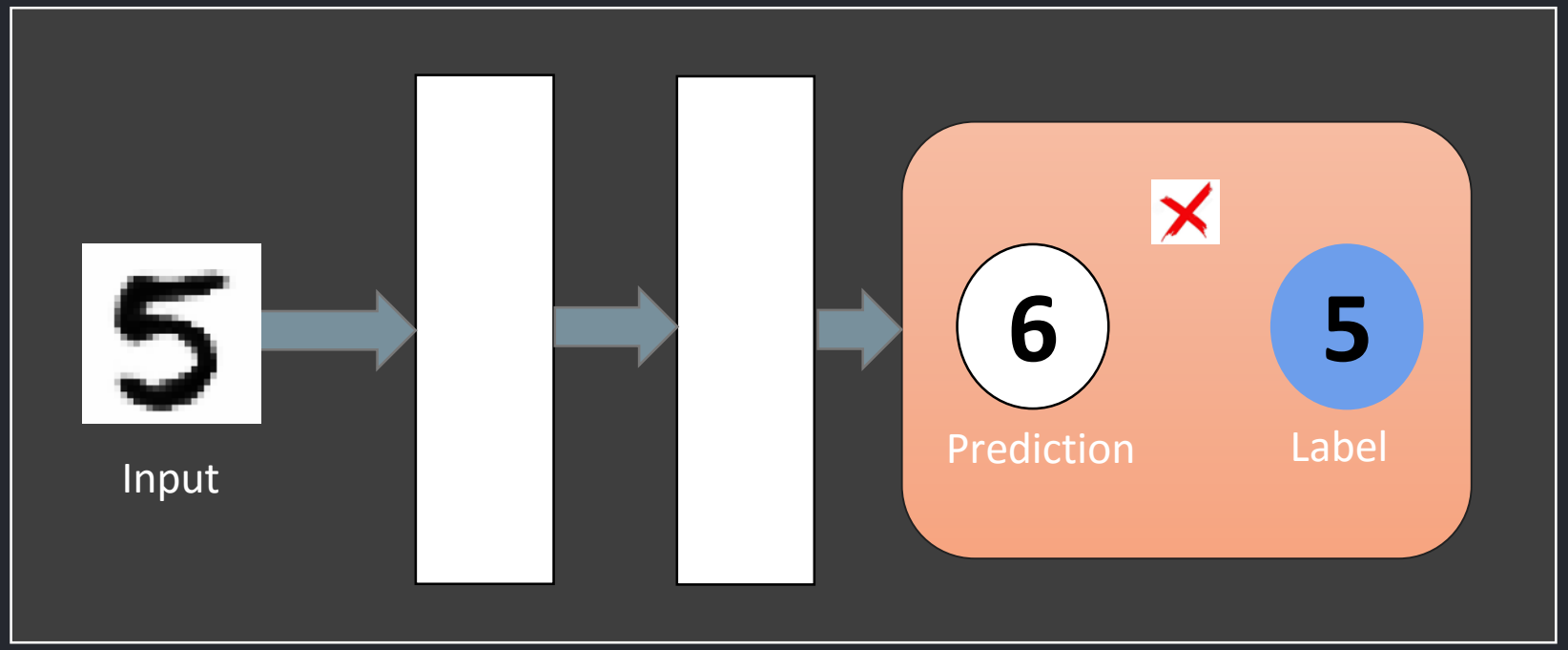
02 Evaluation / Objective

Goal / target of the learning

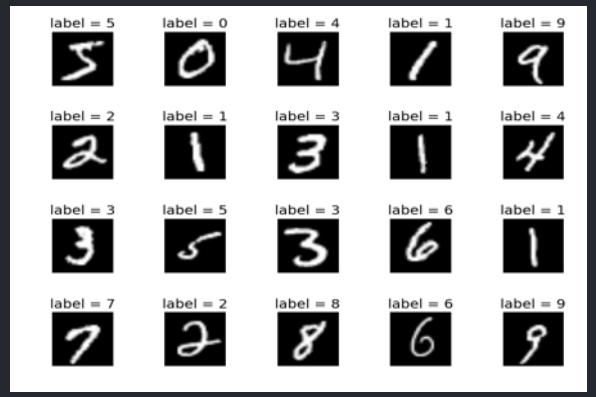
03 Training / Optimization

Learning process to achieve the objective

Learning Illustration

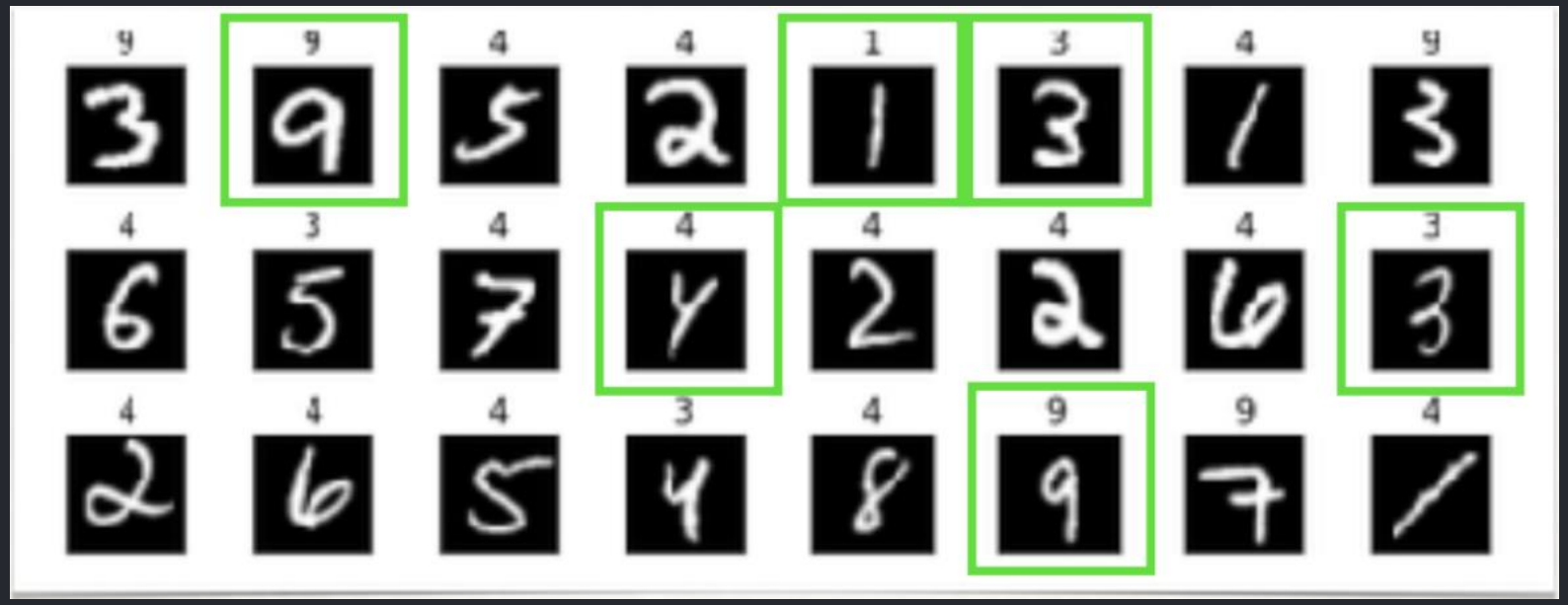


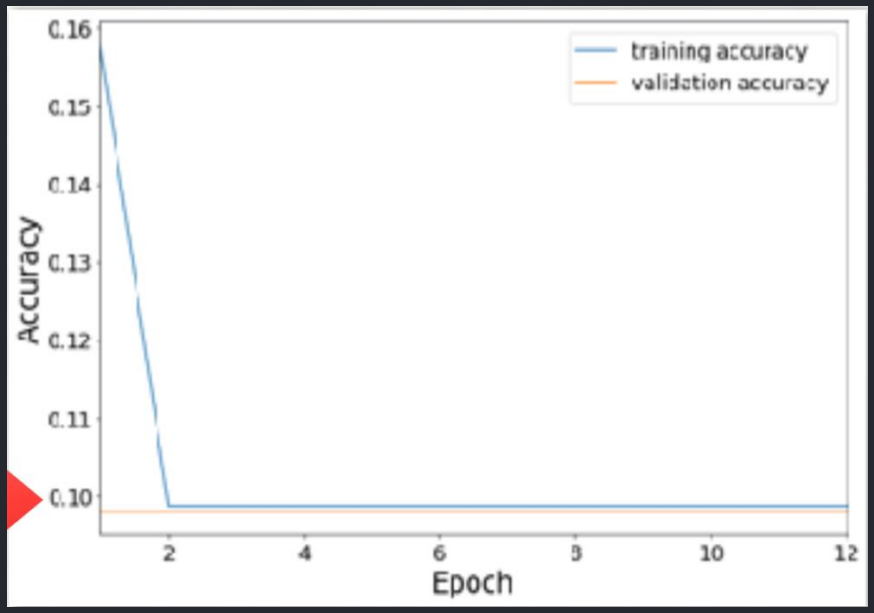
Training data



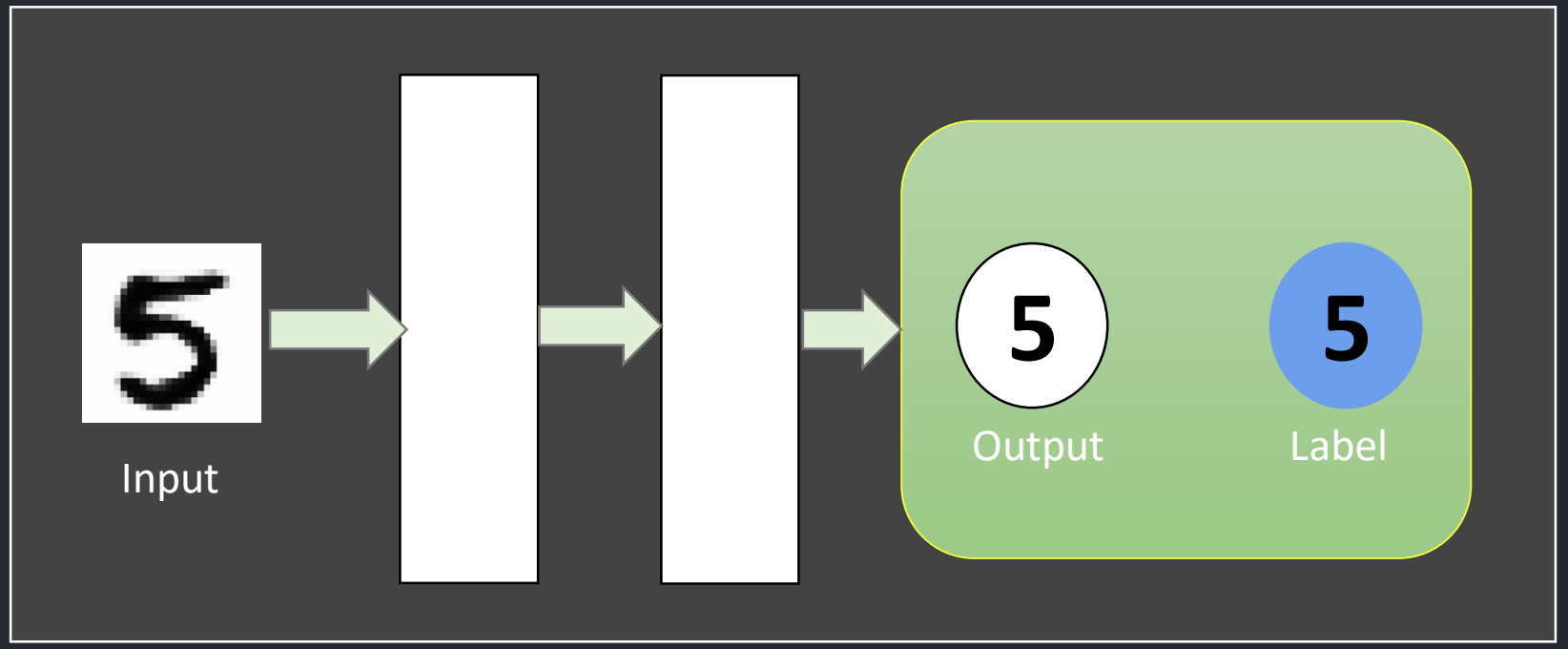
Learning Illustration

Not expected

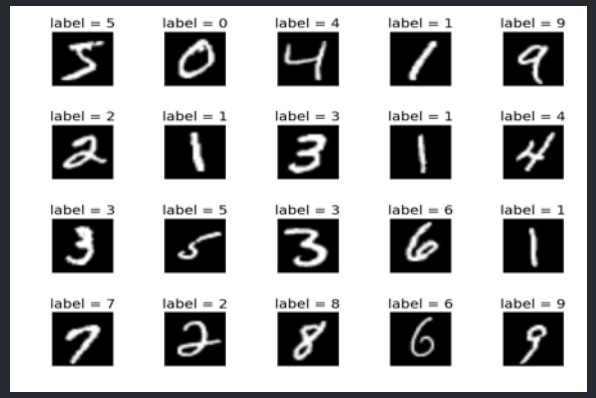




Learning Illustration

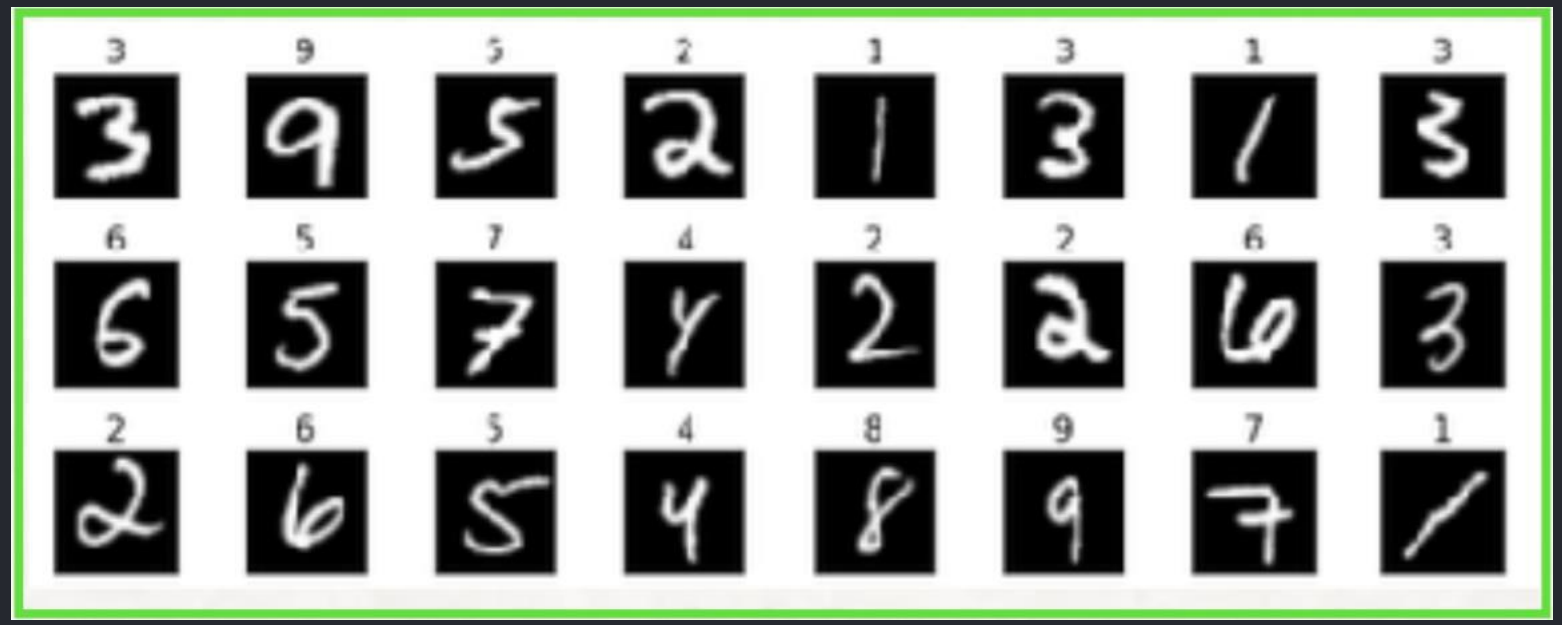


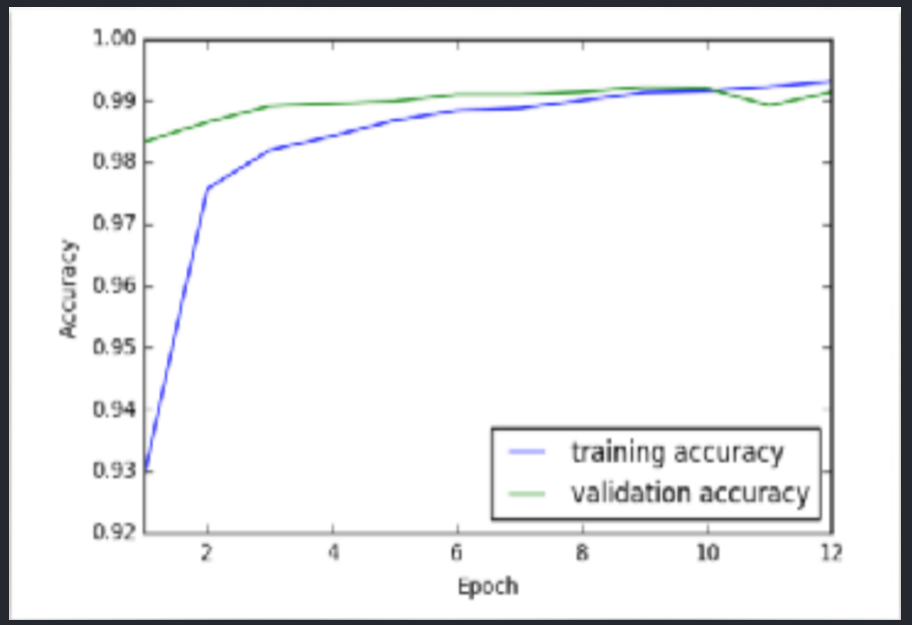
Training data



Learning Illustration

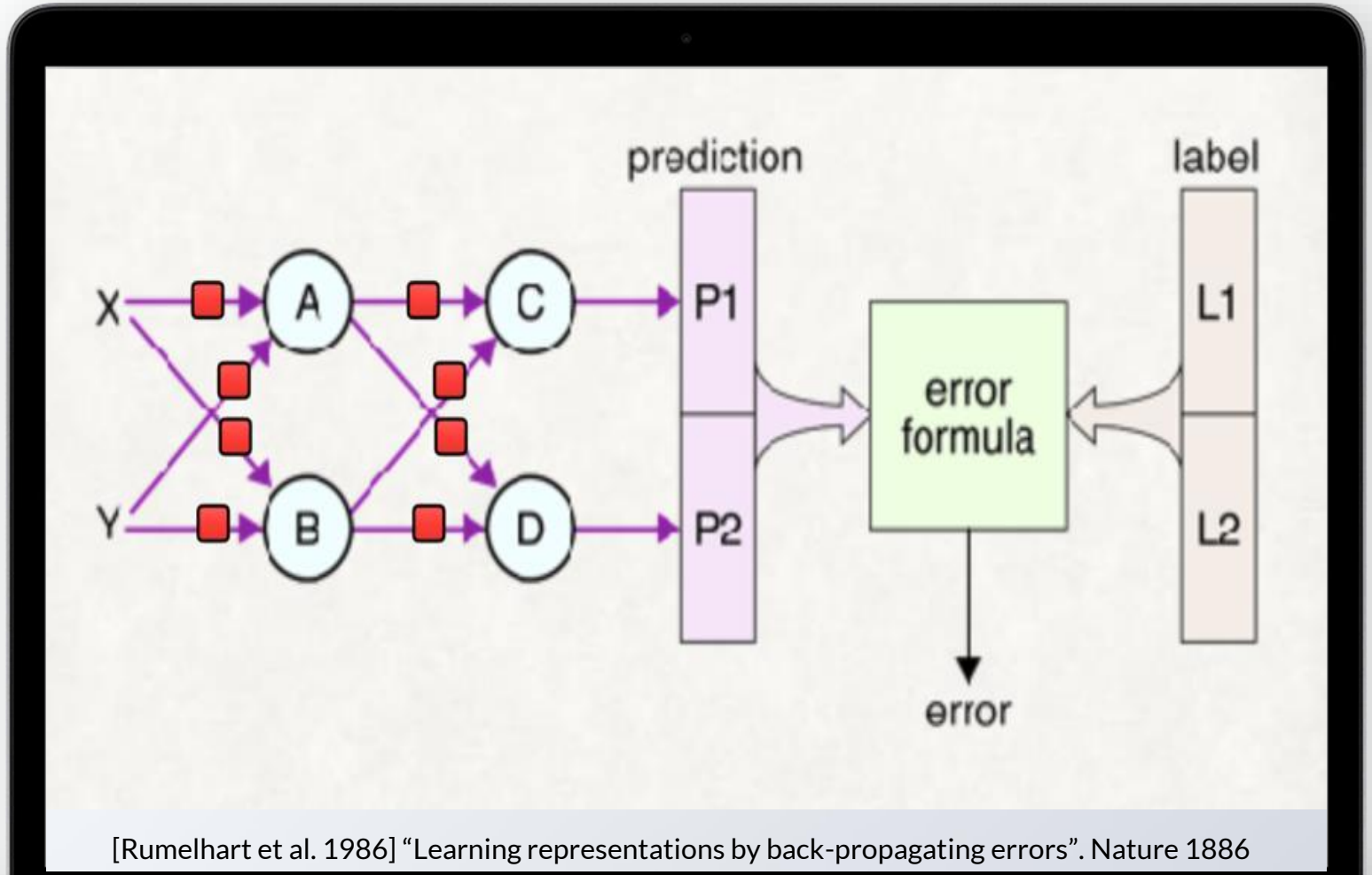
This is what we want!





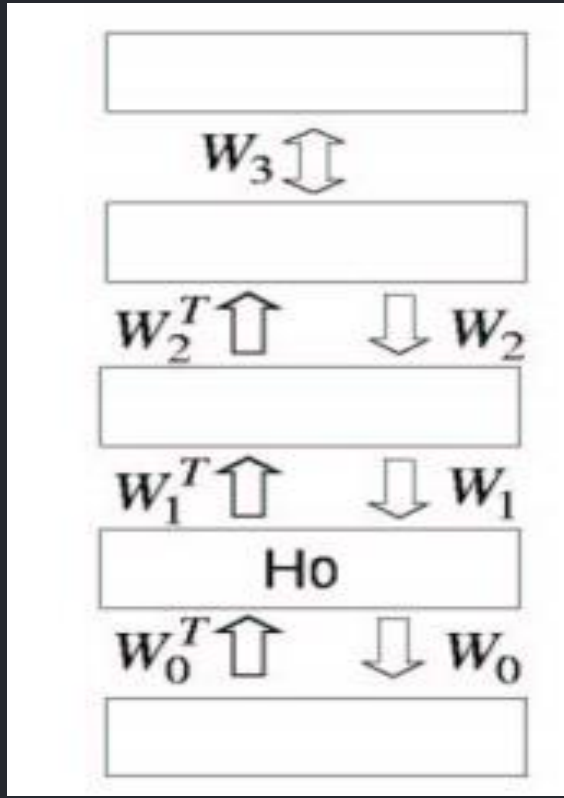
Training / Optimization

Backpropagation



[Rumelhart et al. 1986] “Learning representations by back-propagating errors”. Nature 1886

Where does “Deep Learning” comes from?



The term was first introduced by Rina Dechter (1986), but it wasn’t quite popular, and not exclusive to neural networks.

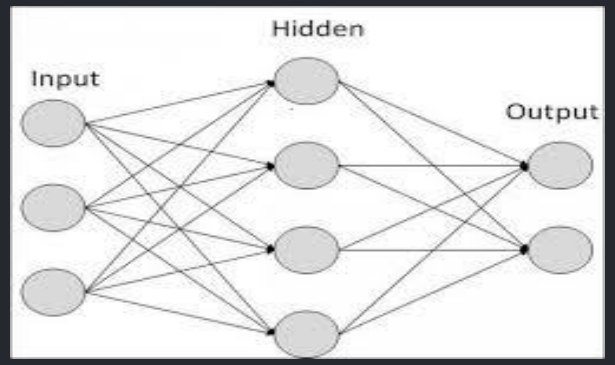
Neural networks with **more than 2 layers** were notoriously under-performed, even compared to those with fewer layers.

In 2006, there were evidences that those, i.e., “deep neural networks”, can outperform standard models, with a particular training strategy.

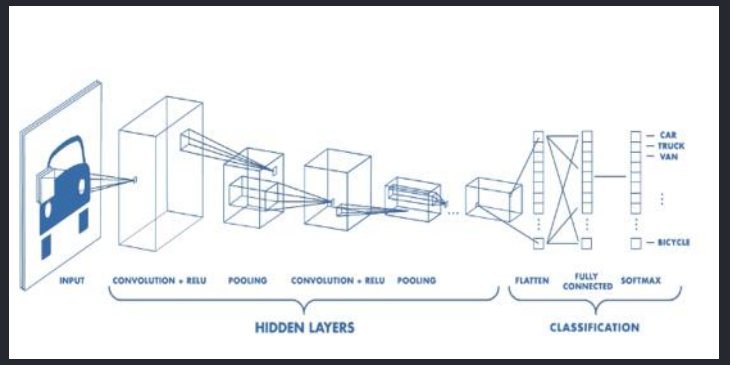
In 2012 onwards, the term has become a new standard to represent the family of neural networks.

[Hinton et al. 2006] “A fast learning algorithm for deep belief nets”. Neural Computation 2006
[Bengio et al. 2007] “Greedy layer-wise training of deep networks”. NIPS 2007

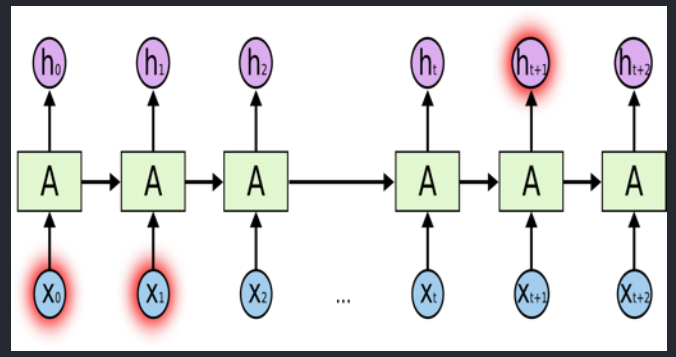
Multi-layer Perceptron (MLP)



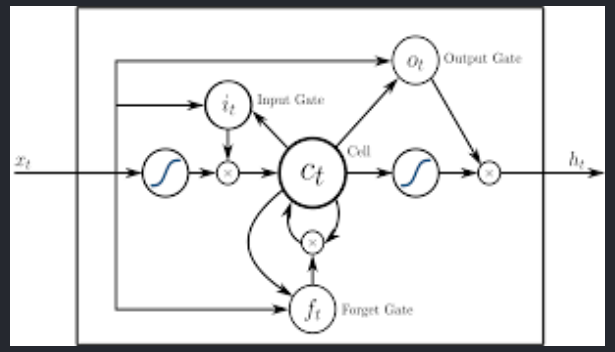
Convolutional Networks (ConvNet)



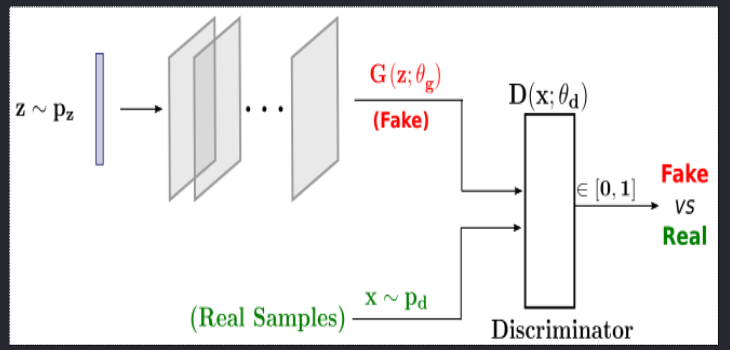
Recurrent Neural Networks (RNN)



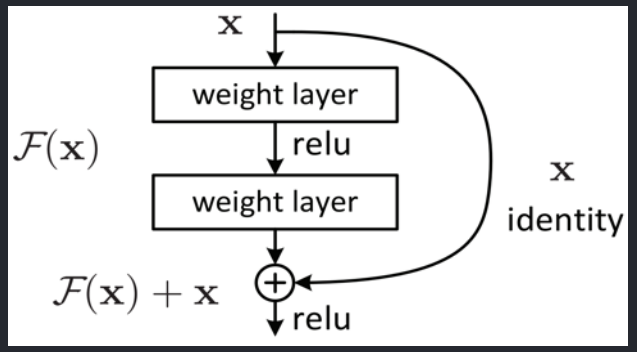
Long-Short Term Memory (LSTM)

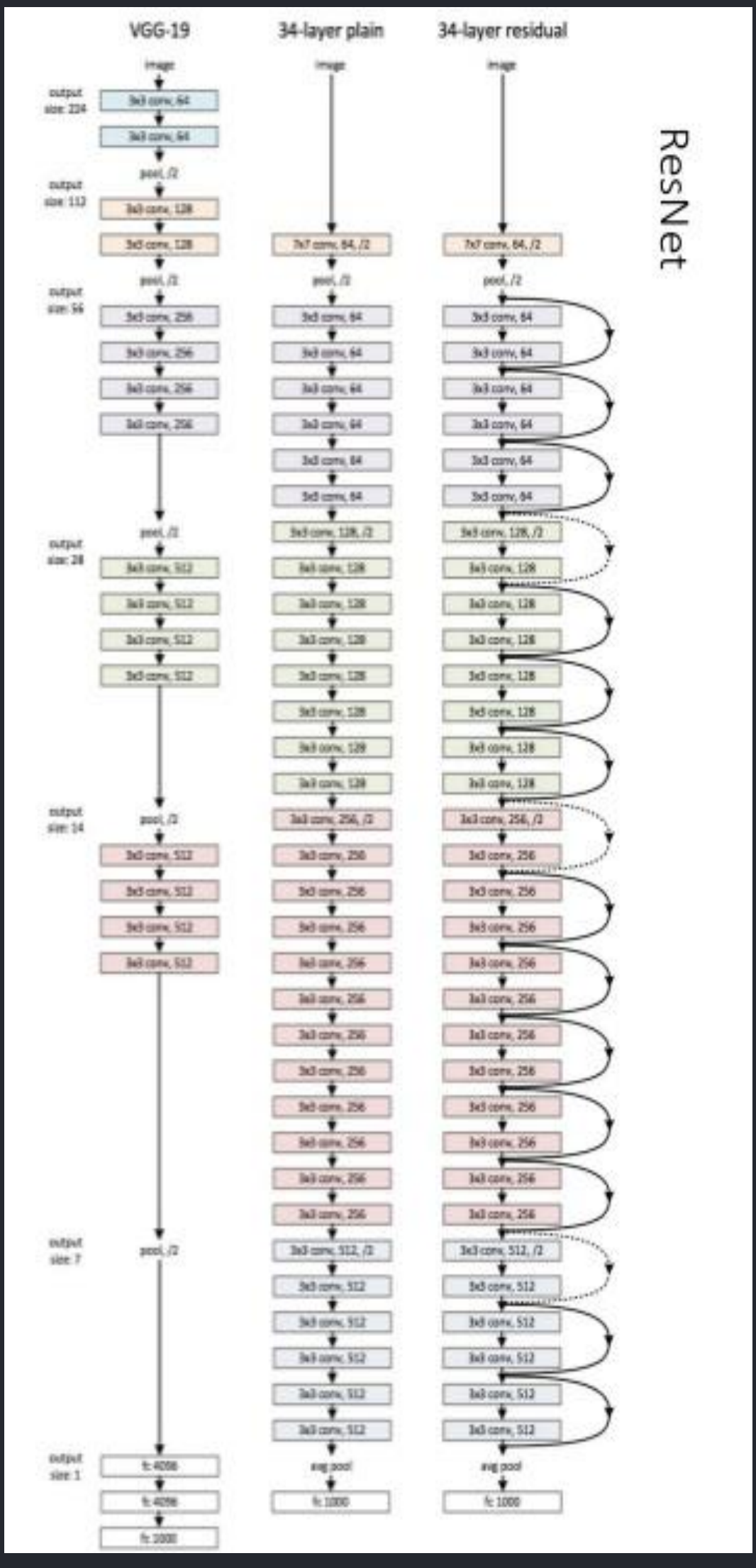


Generative Adversarial Networks (GAN)



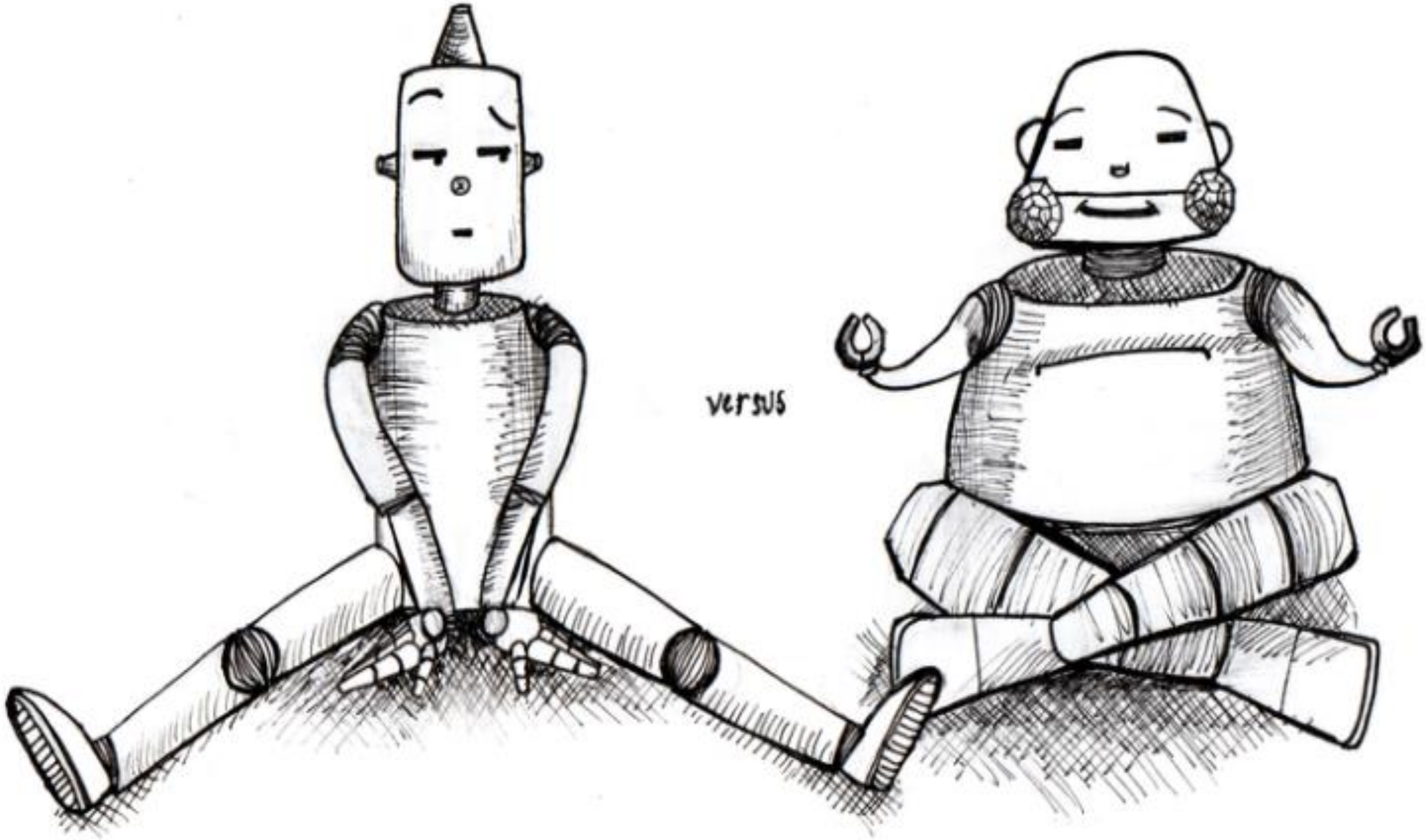
Residual Networks (ResNet)




VGG-19
34-layer plain
34-layer residual
ResNet
image
output size: 224
output size: 112
output size: 56
output size: 28
output size: 14
output size: 7
output size: 1
pool, /2
avg pool
fc 4096
fc 1000

Key Takeaways

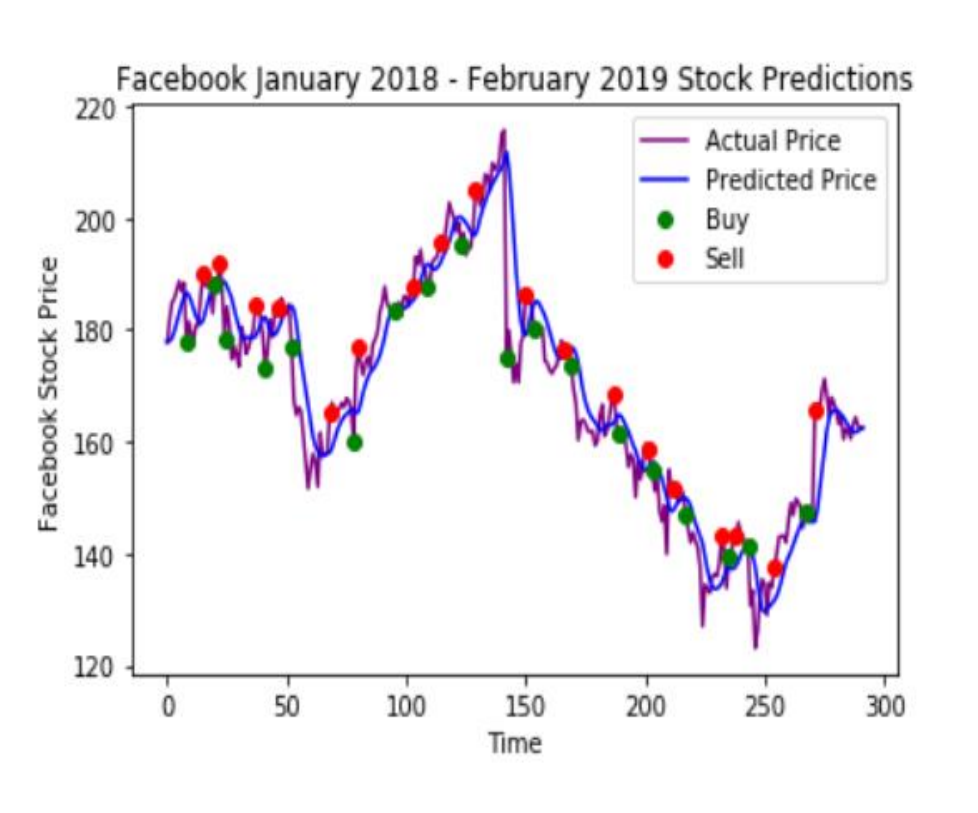
1. **Perceptrons** are the smallest units of deep learning, inspired from neurons in biological brains.
2. A set of perceptrons forms a **neural network**
3. Neural networks learn from mistakes through **back-propagation**
4. **Deep learning** is a new sub-field (with an old history) representing the family of “deep” neural networks

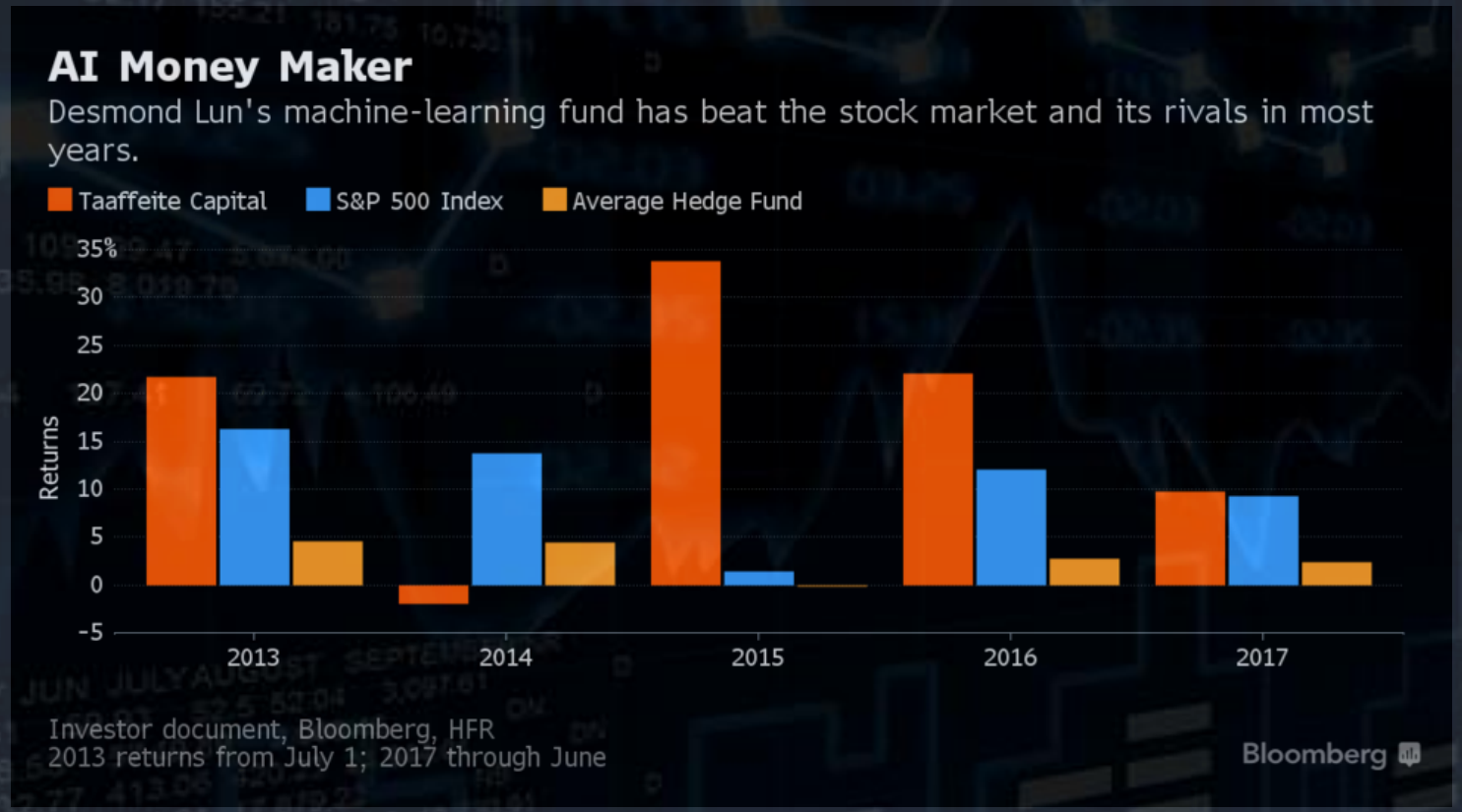

versus
MACHINE LEARNING
DEEP LEARNING

Recent Trend & Issues

Deep Learning

Investment Strategy





Bukalapak

The smaller data, the better

CENTER FOR DATA INNOVATION

ABOUT US PUBLICATIONS BLOG ISSUE REGIONS EVENTS PRESS

Home › Issue › Artificial Intelligence › AI Needs Better Data, Not Just More Data

AI Needs Better Data, Not Just More Data

by Joshua New | March 20, 2019

Harvard Business Review

Innovation | The Future of AI Will Be About Less Data, Not More

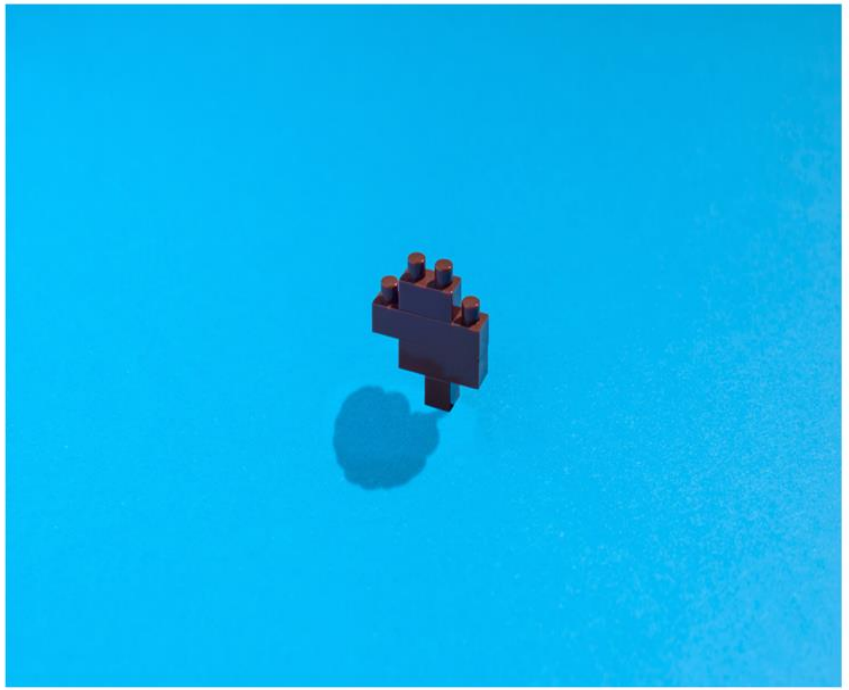
INNOVATION

The Future of AI Will Be About Less Data, Not More

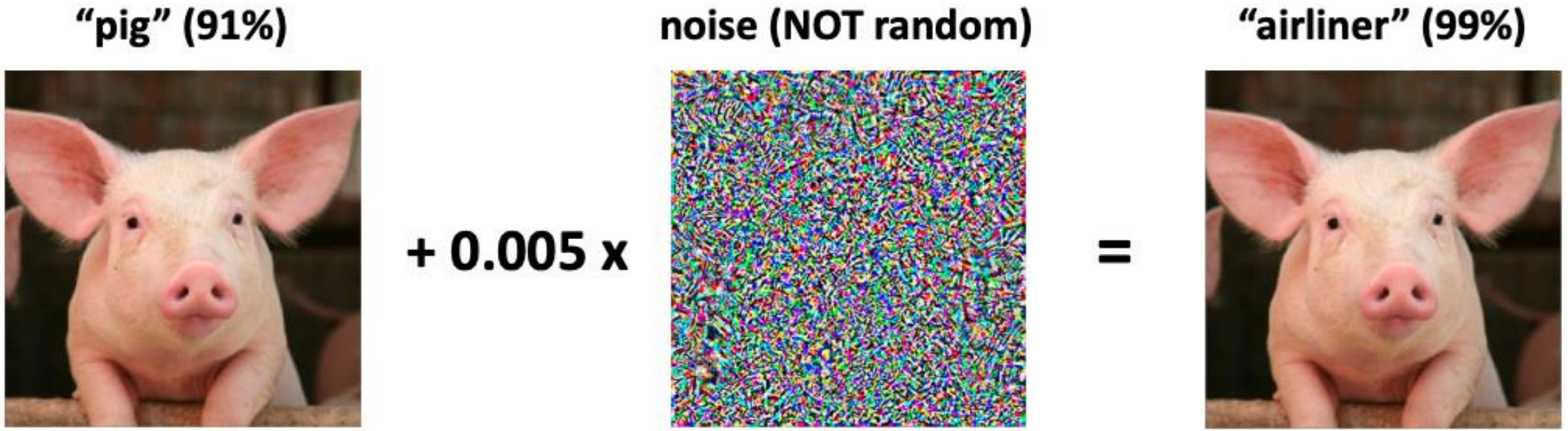
by H. James Wilson, Paul R. Daugherty, and Chase Davenport

JANUARY 14, 2019

Summary Save Share 10 Comment Text Size Print $8.95 Buy Copies



Robustness: Adversarial Attacks



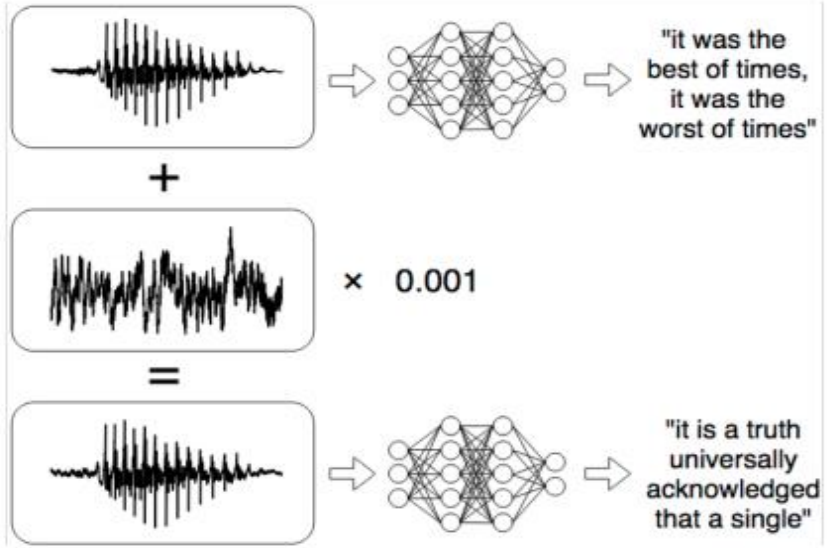
[Szegedy Zaremba Sutskever Bruna Erhan Goodfellow Fergus 2013]
[Biggio Corona Maiorca Nelson Srndic Laskov Giacinto Roli 2013]

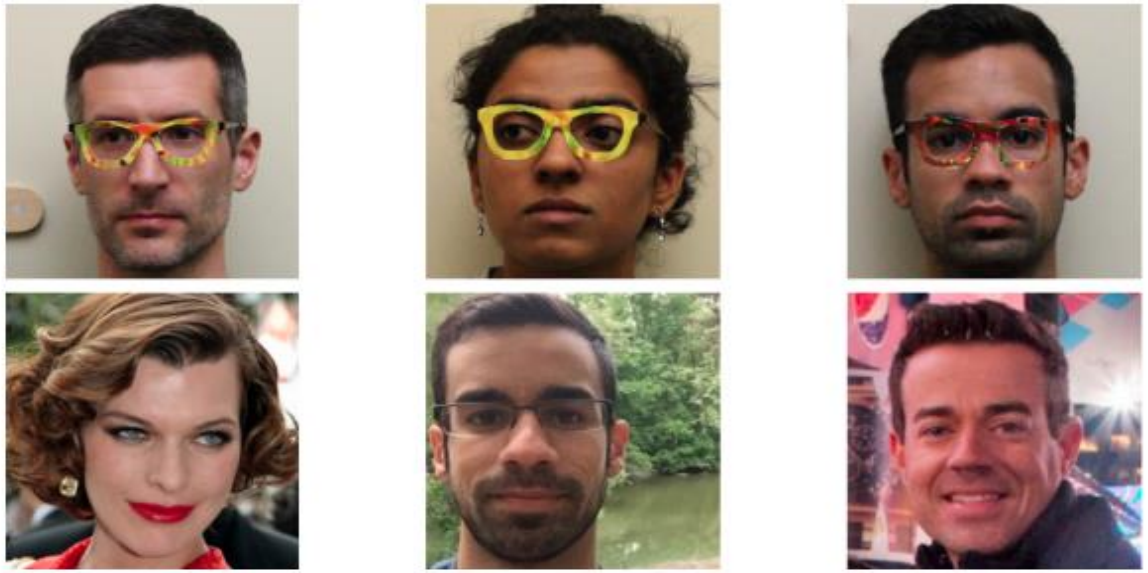
But also: [Dalvi Domingos Mausam Sanghai Verma 2004][Lowd Meek 2005]
[Globerson Roweis 2006][Kolcz Teo 2009][Barreno Nelson Rubinstein Joseph Tygar 2010]
[Biggio Fumera Roli 2010][Biggio Fumera Roli 2014][Srndic Laskov 2013]

Robustness: Adversarial Attacks

→ Security

[Carlini Wagner 2018]:
Voice commands that are unintelligible to humans





[Sharif Bhagavatula Bauer Reiter 2016]:
Glasses that fool face recognition

Fairness

It's Too Late—We've Already Taught AI to Be Racist and Sexist

Now what?

SHARE TWEET

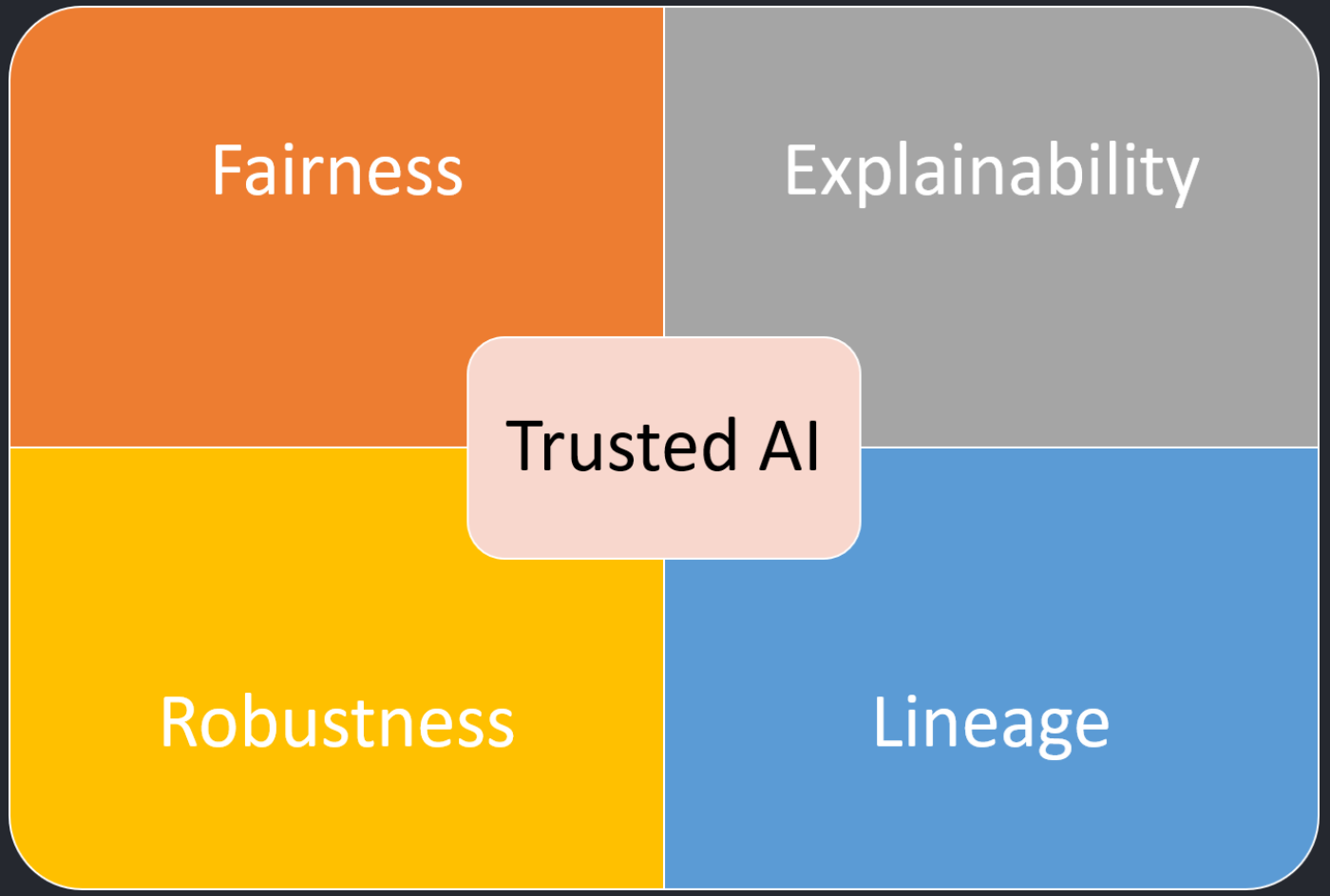
Image: Shutterstock/studiostoks

They say that kids aren't born sexist or racist—hate is taught. Artificial intelligence is the same way, and humans are fabulous teachers.

ProPublica reported, for example, that an algorithm used to to predict the likelihood of convicts committing future crime tends to tag black folks as higher risk than whites. Despite the oft-repeated claim that such data-driven approaches are more objective than past methods of determining the risk of recidivism or anything else, it's clear that our very human biases have rubbed off on our machines.

Consider the case of Microsoft's simple Tay bot, which sucked up all the slurs and racist opinions that Twitter users threw at it and ended up spouting Nazi drivel.

Towards Trusted AI



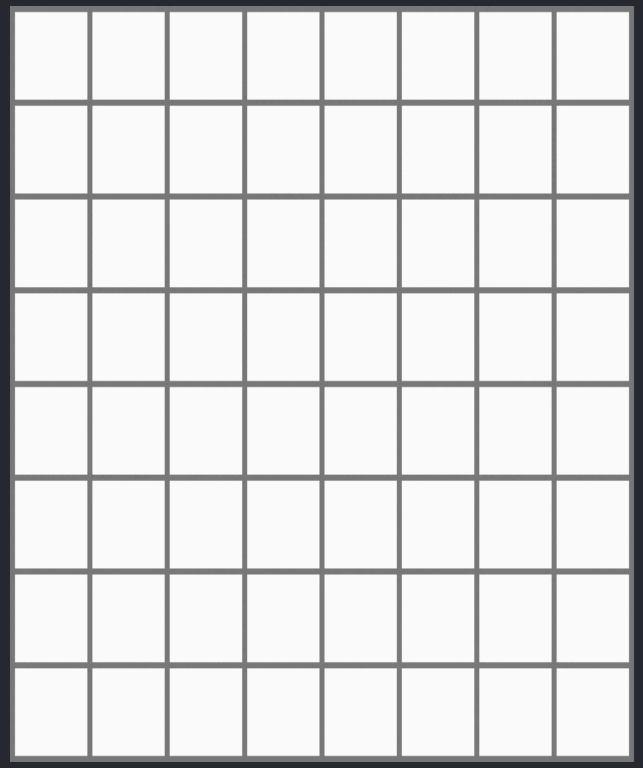
https://towardsdatascience.com/towards-ai-transparency-four-pillars-required-to-build-trust-in-artificial-intelligence-systems-d1c45a1bdd59

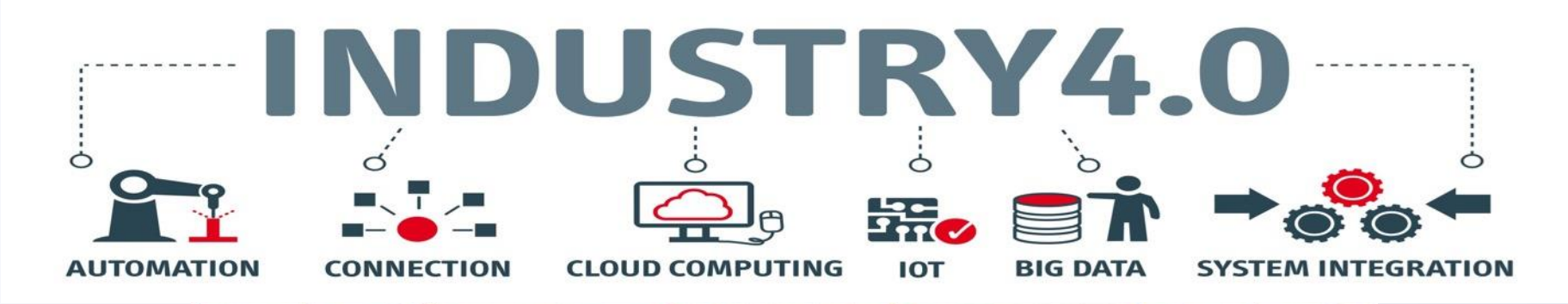
Speech + Video Synthesis

Faking Obama's speech

Art Generator

Learning how to draw





Deep Learning for the Internet of Things

Shuochao Yao and Yiran Zhao, University of Illinois Urbana-Champaign (UIUC)
Aston Zhang, Amazon AI
Shaohan Hu, IBM Thomas J. Watson Research Center
Huajie Shao and Chao Zhang, UIUC
Lu Su, State University of New York, Buffalo
Tarek Abdelzaher, UIUC

How can the advantages of deep learning be brought to the emerging world of embedded IoT devices? The authors discuss several core challenges in embedded and mobile deep learning, as well as recent solutions demonstrating the feasibility of building IoT applications that are powered by effective, efficient, and reliable deep learning models.


I REMEMBER WHEN ONLY A DEEP LEARNING SUPER COMPUTER COULD BEAT ME IN A DATA SCIENCE COMPETITION!
2ND PLACE
1ST PLACE!
COFFEE
KDNUGGETS.com

AI in Bukalapak

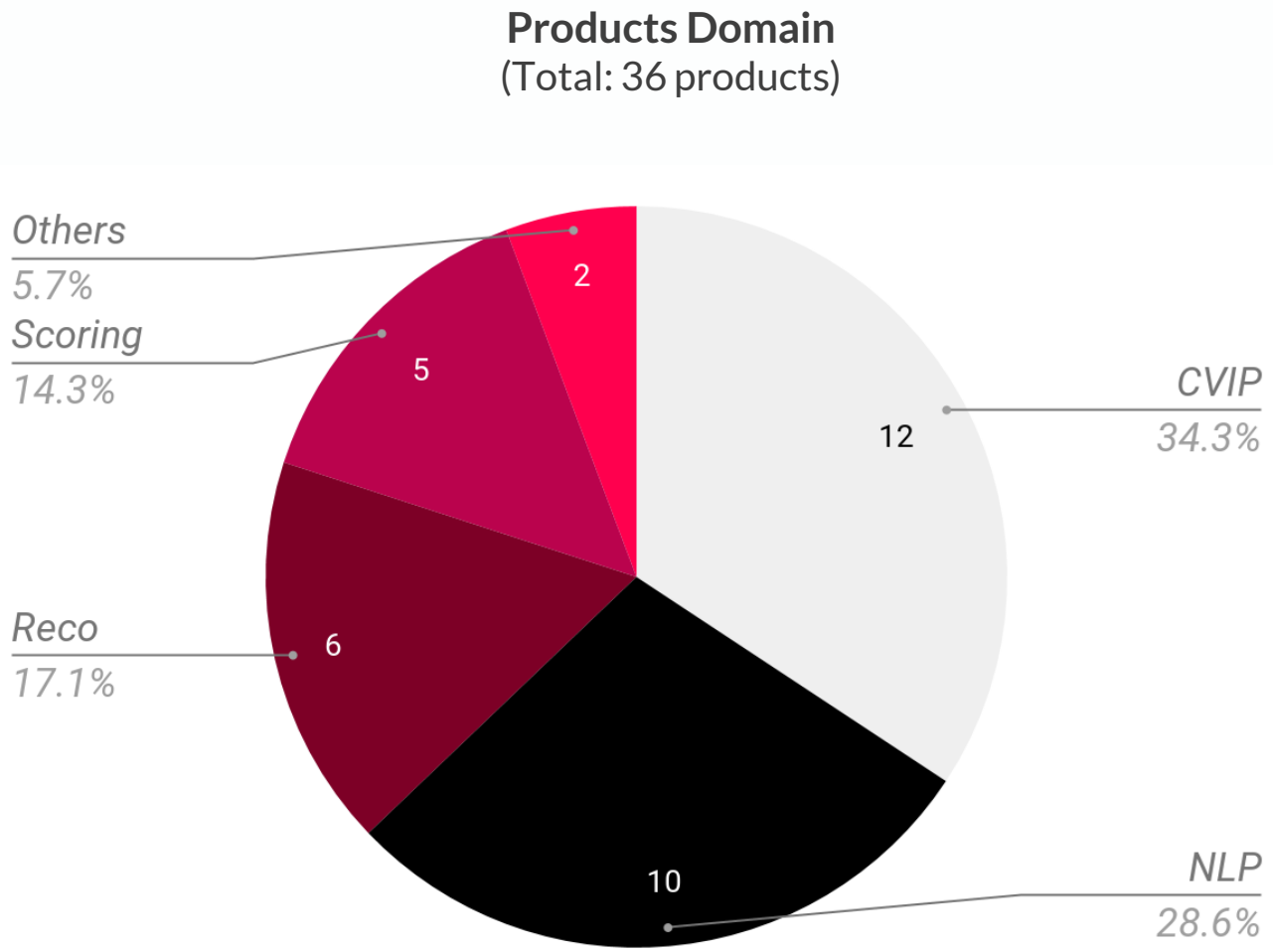
24 June 2019

Artificial Intelligence Research

We are Bukalapak's team that utilizes machine learning (ML) and artificial intelligence (AI) for the company

Currently, we have developed 27 services used across 36 products in Bukalapak. Our domain:

- NLP
- CVIP
- Recommendation
- Scoring



Service

CVIP

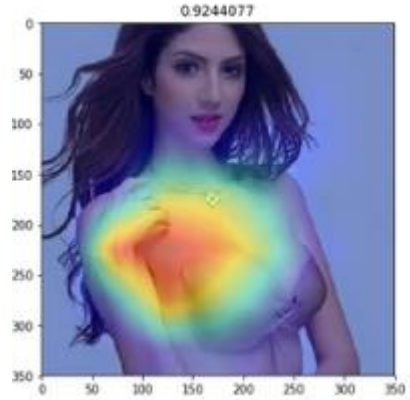
Used in

Sexiness Detection for Online Marketing

Bukalapak use online marketing on Facebook, Google Shopping, and Google Display Network, which will show product pictures to potential users. This service helps score and filter product pictures based on its sexiness. Sexy pictures are considered bad ads because they have low conversion rate, and give Bukalapak a bad name.

Return on Ads Spending

20 %



Service

CVIP

Used in

Logo Detection

Detect and flag competitors logo and delete the image/product since it violates Bukalapak's terms and condition. Gogoist also detect Bukalapak logo from profile pictures.

Flagged items will be reviewed by policy team.

Impact

- Cleaner products
- Decreasing number of reported product
- Decreasing number of reported sellers

Service

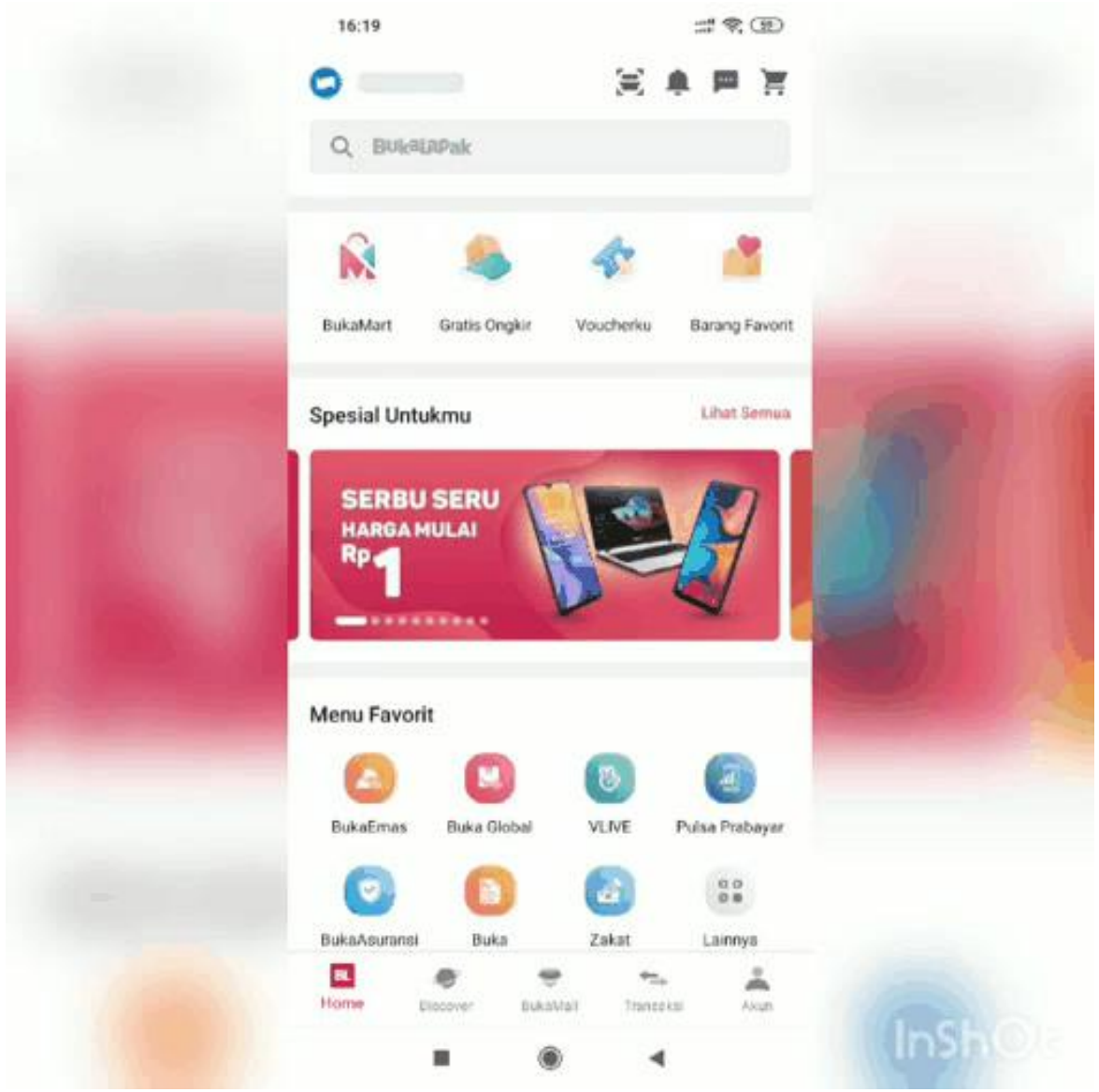
CVIP

Used in

BukaStruk

BukaStruk let users scan their Indomaret and Alfamart bills to get some Credits. Our AI service enables the text extraction and total expense calculation from the scanned receipt automatically.

A user will get Rp 200 as credits for each successful scan.



Service

NLP

Used in

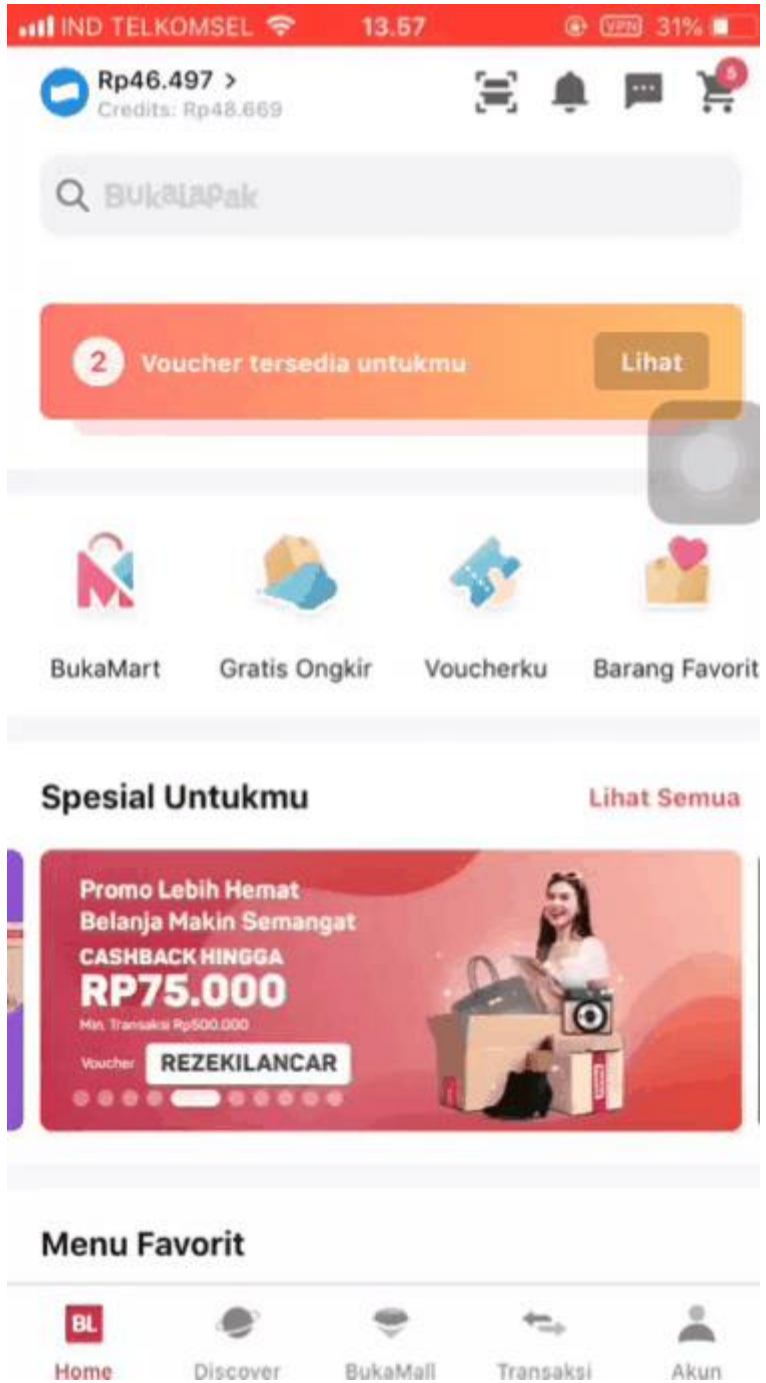
Catalog

This service helps the automation of grouping of the products in Catalog.

Product category using baymax:

- Handphone
- Laptops
- Camera

By cataloging product, we expect it will make user browsing similar products instead search it one by one.



Service

Recommendation

Used in

Recommendation on PDP, Cart, and Homepage

Providing relevant recommendation based on personalization, popular products, and sellers to users. Recommendation is in Product Detail Page, Cart, and Homepage

Usage period

Since 2017

Total GMV Attribution

> Rp 1 trillion

Further Reading:

- Recommendation impact in Bukalapak

Service

Scoring

Used in

BukaModal

Credit scoring assigns scores used to identify the risk of default of the debtor, based on character, capacity to pay, and ability to pay.

Revenue is fee which Bukalapak will get for every disbursement based on agreement with partners.

Potential GMV is the amount of disbursement to users chosen by the service.

Potential GMV

> Rp 5 Bio

Service

Scoring

Used in

Saldo Bantuan Mitra

Credit scoring assigns scores used to identify the risk of default of Mitras. Prediction is based on character, capacity to pay, and ability to pay.

Loss prevention is the amount of loss prevented because of better Mitra scoring

Loss Prevention

> Rp 200 Mio

Service

Scoring

Used in

IHSG and Gold Price Notifications

We alert investors with custom notifications that match with users investment profile. For a start, we will implement general notifications about gold price and IHSG.

GMV Contribution

~ 20 % of GMV

Bukareksa / Bukaemas

Thank you

Deep Learning: A Breakthrough in AI

Muhammad Ghifary, PhD

Head of Artificial Intelligence

muhammad.ghifary@bukalapak.com

Joelian Samuel, M.Sc.

AI Research Manager

joelian.samuel@bukalapak.com