Tổng quan về Trí Tuệ Nhân Tạo

TS. Đỗ Như Tài
Đại Học Sài Gòn
dntai@sgu.edu.vn

AI Trong Cuộc Cách Mạng Công Nghệ Lần Thứ 4

3

4

Definition

of

Artificial Intelligence (A.I.)

Artificial Intelligence

"… the **science** and

engineering

of

making

intelligentmachines”

(John McCarthy, 1955)

Source: https://digitalintelligencetoday.com/artificial-intelligence-defined-useful-list-of-popular-definitions-from-business-and-science/

Artificial Intelligence

"... technology that thinks and

acts

like humans"

Source: https://digitalintelligencetoday.com/artificial-intelligence-defined-useful-list-of-popular-definitions-from-business-and-science/

Artificial Intelligence

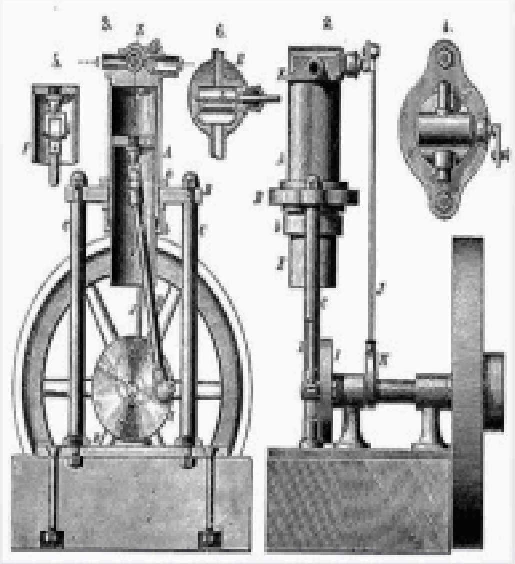
"… intelligence exhibited by

machines or software”

Source: https://digitalintelligencetoday.com/artificial-intelligence-defined-useful-list-of-popular-definitions-from-business-and-science/

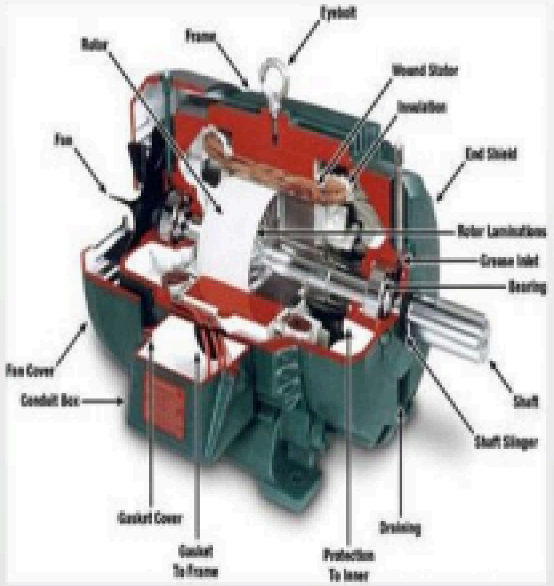
AI là Động cơ lõi (core engine) của CMCN 4.0

Lần thứ nhất



Sản xuất cơ khí với máy móc dựa vào động cơ hơi nước (Cuối thế kỷ 18, đầu tk 19)

Lần thứ hai



Sản xuất hàng loạt với máy móc dựa vào năng lượng điện (Cuối tk 19, đầu tk 20)

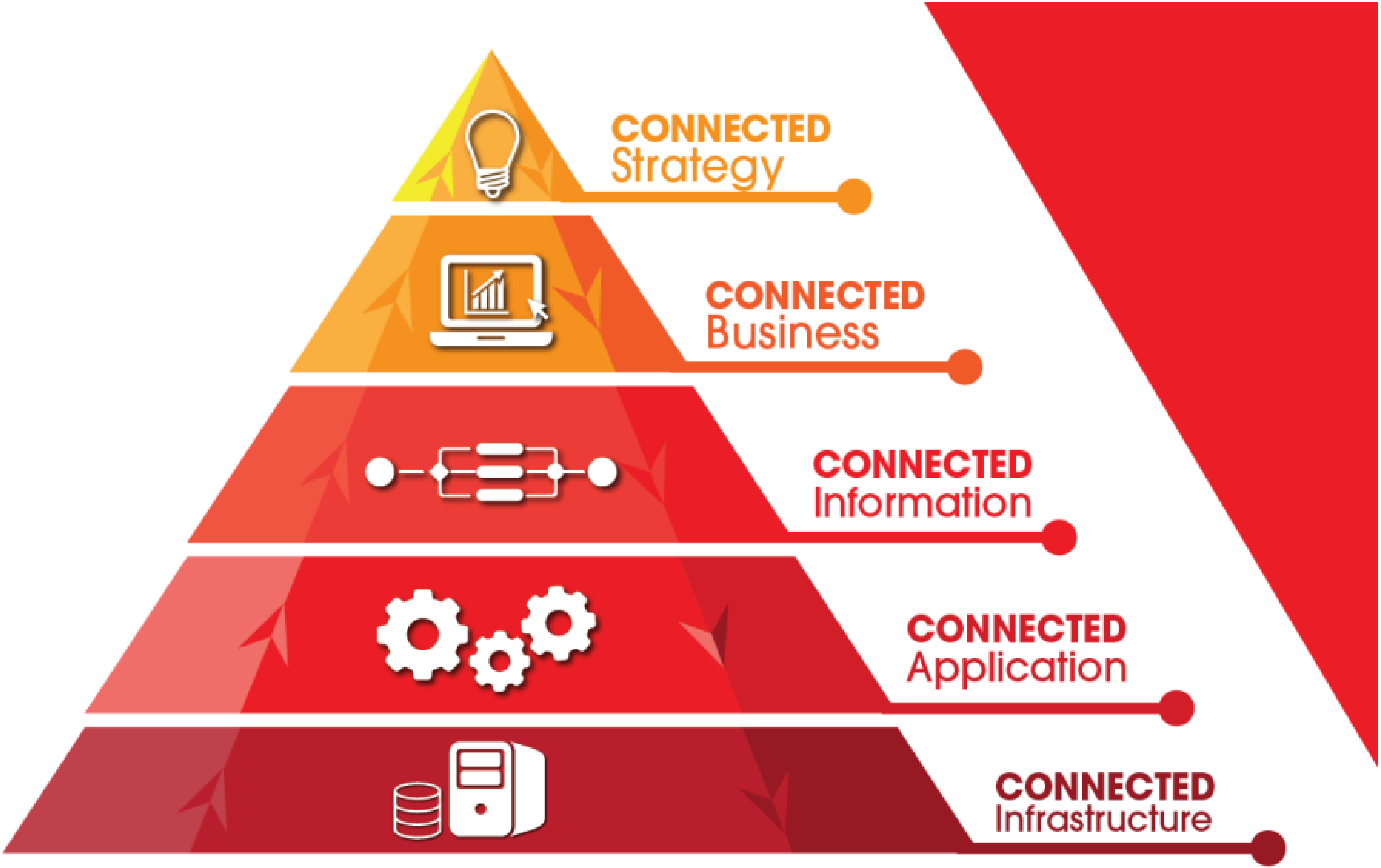
Lần thứ ba

Sản xuất tự động với máy tính, điện tử và cách mạng số hóa (Từ 1970)

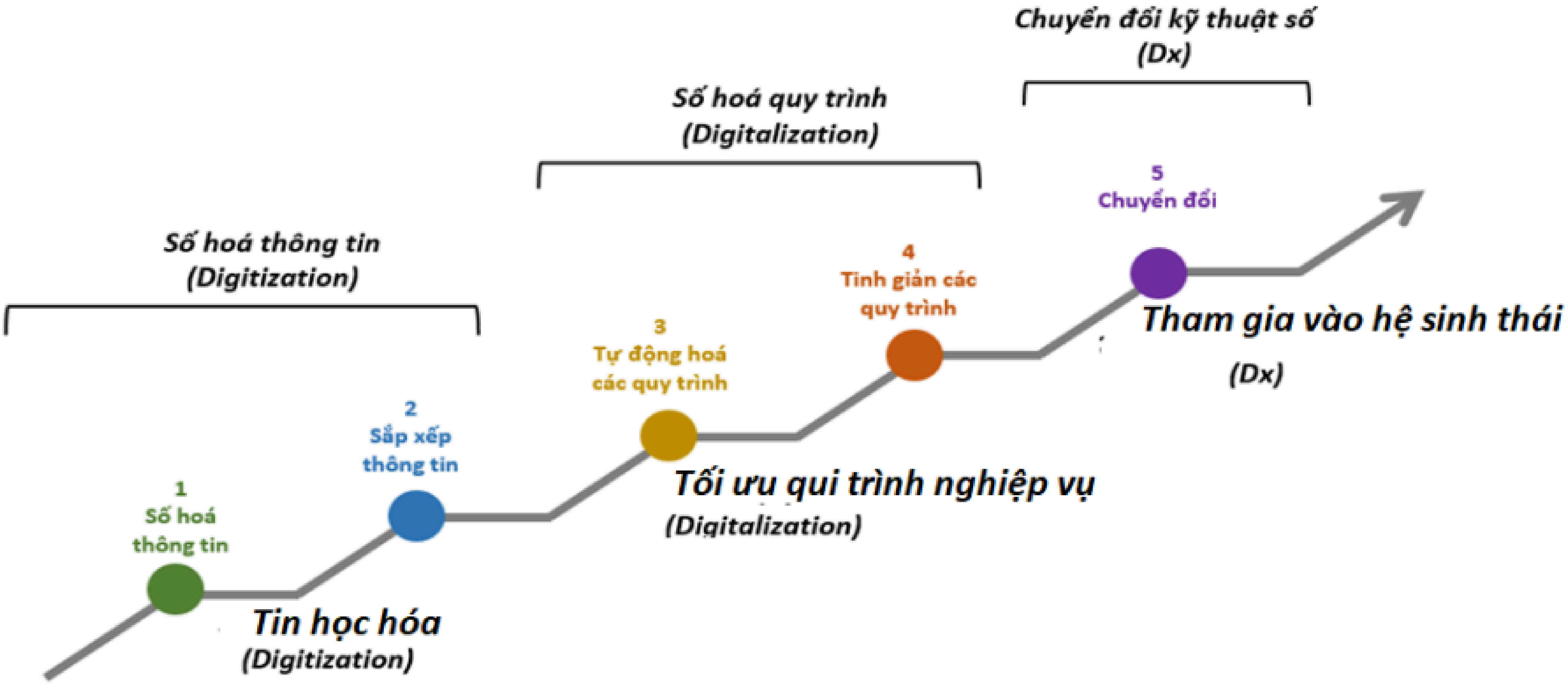
Lần thứ Tư

Sản xuât thông minh nhờ chuyển đổi số - kinh tế chia sẻ (Từ bây giờ, đúng thời điểm)

Chuyển đổi số



Lộ trình chuyển đổi số



Nền Kinh Tế Chia Sẽ


UBER
lyft
BREEZE
wingz
JustPark
airbnb
HomeAway
wework
amazon mechanical turk
Artificial Artificial Intelligence
TaskRabbit
freelancer
ebay
craigslist
Etsy

Câu hỏi

- AI là gì?

- CMCN lần 4.0 vs AI vs Chuyển đổi số?
- Mục tiêu của chuyển đổi số?
- Lộ trình chuyển đổi số?
- Là lực lượng sản xuất mới, anh (chị) làm gì để thích ứng với quan hệ sản xuất mới và phương thức sản xuất mới?

Deep in Dive AI

AI, Big Data, Cloud Computing

Evolution of Decision Support, Business Intelligence, and Analytics

AI

AI Cloud Computing Big Data DM BI

Source: Ramesh Sharda, Dursun Delen, and Efraim Turban (2017),
Business Intelligence, Analytics, and Data Science: A Managerial Perspective, 4th Edition, Pearson

4 Approaches of AI

Thinking Rationall y

Thinking Humanly

Acting Humanly

Acting Rationall y

Source: Stuart Russell and Peter Norvig (2016) , Artificial Intelligence: A Modern Approach, 3rd Edition, Pearson International

4 Approaches of AI

2. Thinking Humanly: The Cognitive Modeling Approach

1. Acting Humanly: The Turing Test Approach (1950)

3. Thinking Rationally: The "Laws of Thought" Approach

4. Acting Rationally: The Rational Agent Approach

Source: Stuart Russell and Peter Norvig (2016) , Artificial Intelligence: A Modern Approach, 3rd Edition, Pearson International

AI Acting Humanly: The Turing Test Approach (Alan Turing, 1950)

- Natural Language Processing (NLP)
- Knowledge Representation
- Automated Reasoning
- Machine Learning (ML)
- Computer Vision
- Robotics

Source: Stuart Russell and Peter Norvig (2016) , Artificial Intelligence: A Modern Approach, 3rd Edition, Pearson International

Artificial Intelligence

~~Machine Learning & Deep Learning~~

Source: https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/

Artificial Intelligence (A.I.) Timeline

Source: https://digitalintelligencetoday.com/artificial-intelligence-timeline-infographic-from-eliza-to-tay-and-beyond/

How to define Artificial Intelligence?

Lesson: People are easy to fool!

Lesson: Some Things are Particularly Hard to Code

Use Rules? Use Data?

What is Learning From Data?

Machine Learning

Supervised Learning (Classification)

Learning from Examples

$$y = f(x)$$

x y

input *Output*

label

What is Learning From Data?

Machine Learning
Supervised Learning (Classification)
Learning from Examples

$$y = f(x)$$

Example

5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa

x y

7.0,3.2,4.7,1.4,Iris-versicolor
6.4,3.2,4.5,1.5,Iris-versicolor
6.9,3.1,4.9,1.5,Iris-versicolor
6.3,3.3,6.0,2.5,Iris-virginica

```
5.8,2.7,5.1,1.9,Iris-virginica
7.1,3.0,5.9,2.1,Iris-virginica
```

Time Series Data


```
[10, 20, 30, 40, 50, 60, 70, 80, 90]
```

X Y

```
[10 20 30] 40
[20 30 40] 50
[30 40 50] 60
[40 50 60] 70
[50 60 70] 80
[60 70 80] 90
```


What is Learning From Data?

The Theory of Learning

- How can we be sure that our learned hypothesis will predict well for previously unseen inputs?

- How do we know that the hypothesis h is close to the target function f if we don't know what is?

- How many examples do we need to get a good h?
- What hypothesis space should we use?
- If the hypothesis space is very complex, can we even find the best h or do we have to settle for a local maximum?
- How complex should h be?
- How do we avoid overfitting?

Linear Regression Weight Space

$$h_w(x) = w_1 x + w_0$$

$w^* = \mathbf{argmin}_w \; Loss(h_w)$ ***Loss function for Weights*** $(w_1,$

w_0 $y = 0.232\,x + 246$)

Source: Stuart Russell and Peter Norvig (2020), Artificial Intelligence: A Modern Approach, 4th Edition, Pearson

Linear Regression Weight Space

How to find the best values for w1, w2 and w3 ?

How to find the best values for real-data?

Conclusion?

More data and

computing power
always beat
fancy algorithms

AI From Simple to Complex Models

Perceptron – Architecture

- Biological neuron vs Artificial neuron

Perceptron – Architecture

- Architecture

1 ��

��$_1$

��$_2$

…

��$_{��}$

- input: x=[��$_1$, ��$_2$, … , ��$_{��}$]$^{��}$
- weights: w=[��$_1$, ��$_2$, … , ��$_{��}$]$^{��}$
- bias: ��
- activation function: ��

��$_1$ ��$_2$

��$_{��}$

��

��

- output: �� = �� w^{T}x + �� = �� ��▌��▌+ ��▌��▌+ ⋯ + ��█��█+ ��

Perceptron - Capability

- Single Perceptron with ��_{��} represents a linear classifier in ��^{��}
  - �� = ��_{��}(w^T x + ��) = { +1, w^T x + �� ≥ 0 ; −1, w^T x + �� < 0 }

1 ��

1 −3

��_1 ��_2 ��_{��}

1.5 ��

1.0

��_2

(��_2 = −1.5��_1 + 3) ��_1

��_3 ��_{��}

��_1 ��_2

1.5��_1 + ��_3 ��_2 − 3 = 0

��_1 ��_2

��

��_1��_1 + ��_2��_2 + ��_3��_3 + �� = 0

MLP – Architecture

- Multi-Layer Perceptron
  - input layer, hidden layer(s), output layer
  - feed-forward
  - fully-connected

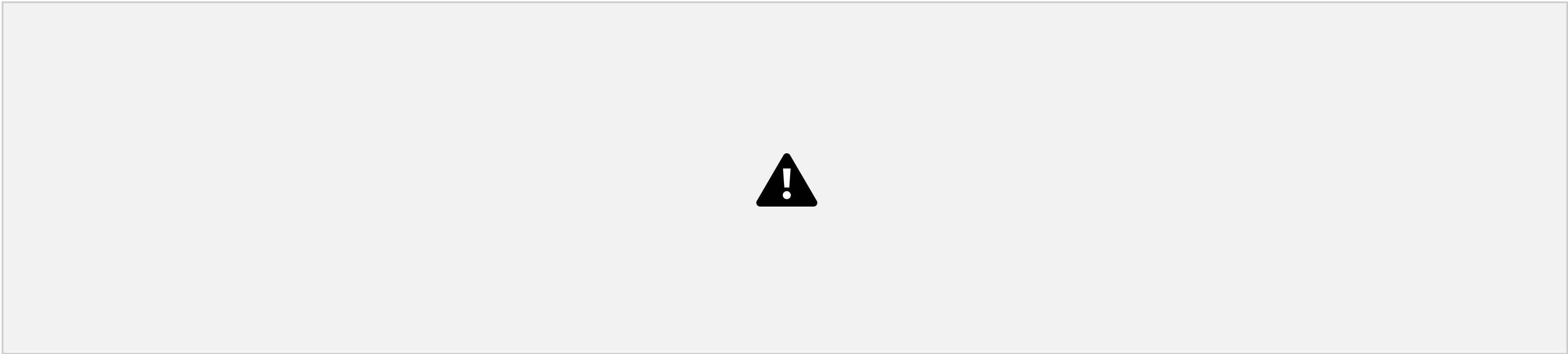
3×3×1 MLP 3×3×2×2 MLP

MLP – Capability

- Nonlinear decision boundaries
  - 2×4×1 MLP with sigmoid activation
  - learning rate = 0.01
  - max #-iterations = 9,000

MLP – Learning: Deep MLP

- Deep MLP



layer 1 ���������� �� ���������� ��

- ��$_{��}$: ��-th input value of layer 1
- ��$_{����}^{(��)}$: weight from ��-th node of layer �� to ��-th node of layer �� + 1
- ��$_{��}^{(��)}$: output from ��-th node of layer ��
- ��$_{��}$(��$_{��}$) : ��-th output(desired output) value in layer ��

CNN - Architecture

- Deep feed-forward network
  - Convolution layers
  - Pooling layers
  - Fully-Connected layers

CNN Architecture

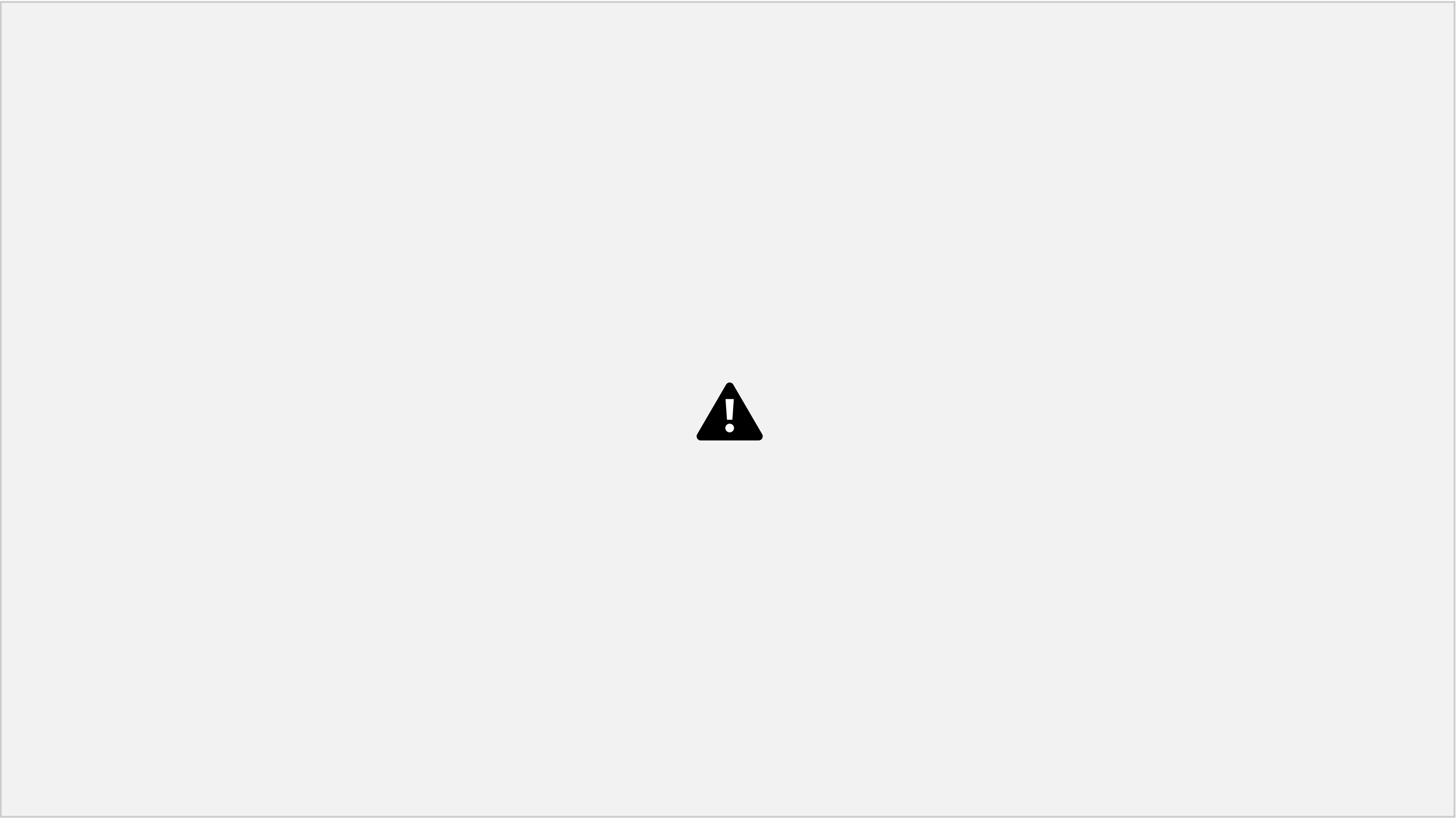
- **Example CNN (AlexNet, 2012)**
  - **Conv:1-Pool:1-Conv:2-Pool:2-Conv:3-Conv:4-Conv:5-Pool:4-FC:6-FC:7-FC:8**

CNN - Architecture

- Feature Maps - multiple levels

CNN - Learning

- Back-propagation algorithm for learning CNN

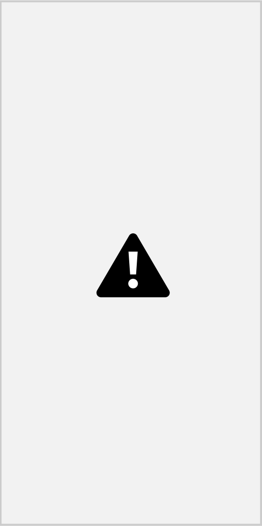


CNN - Capability

- AlphaGo [Nature, 2016]

RNN – Architecture

- Vanilla RNN

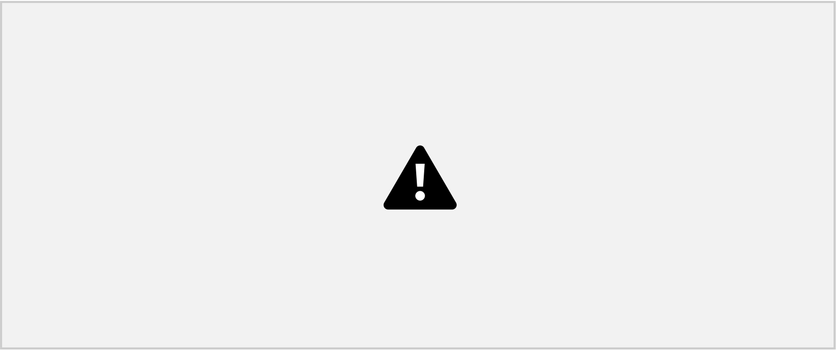
• input sequence: ��1, ��2, … , ����, … , ����
• output sequence: ��1, ��2, … , ����, … , ���� • states: ��0, ��1, ��2, … , ����, … , ���� • weights: ��, ��, ��

���� = tanh ���� [redacted] + ���� [redacted]

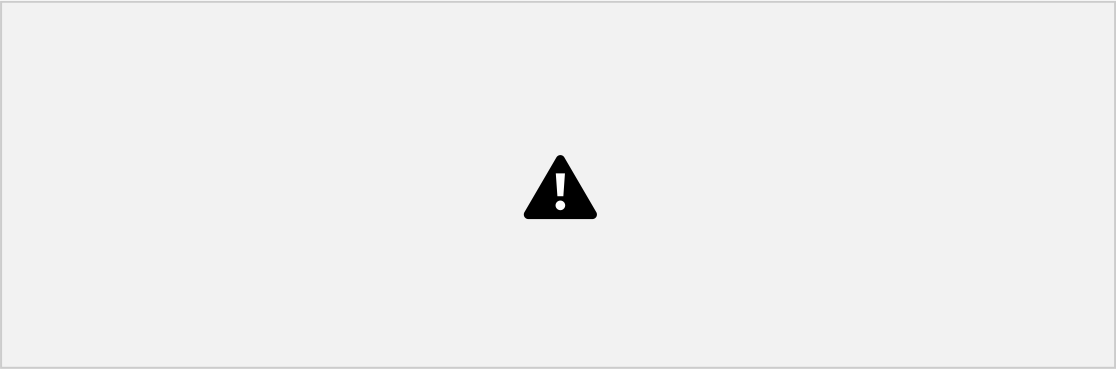
���� = softmax(������)

Useful for Temporal Dependencies

- Short-term dependency

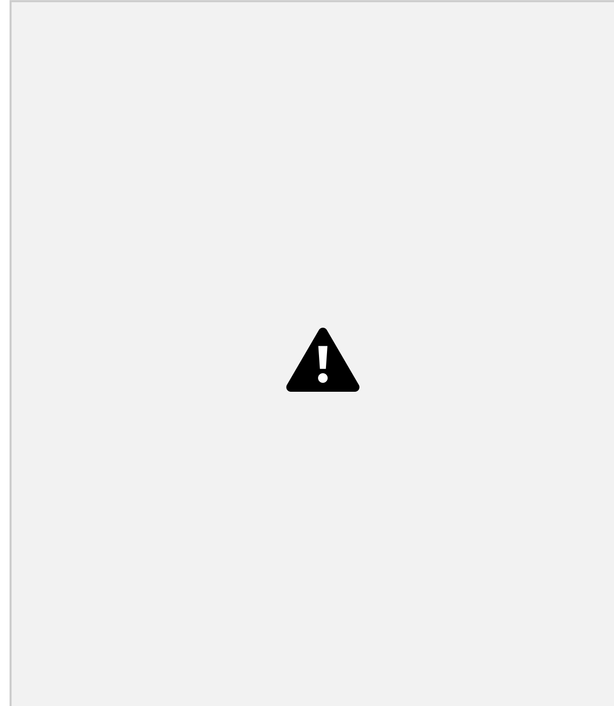


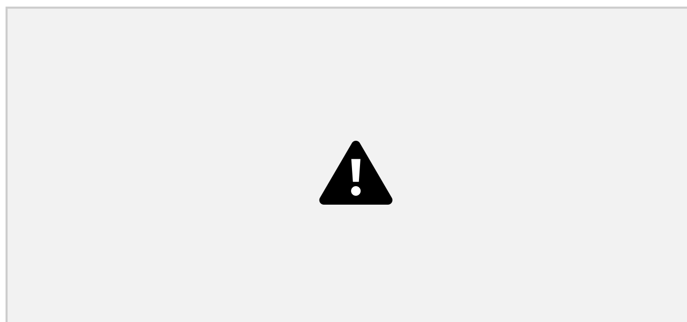
- Long-term dependency



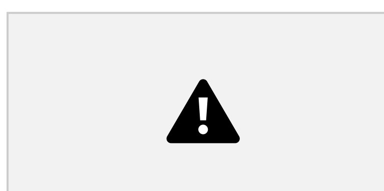
RNN - Applications

- many-to-many
  - Text Summarization

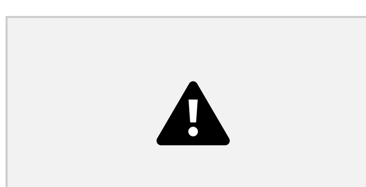


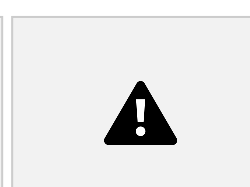
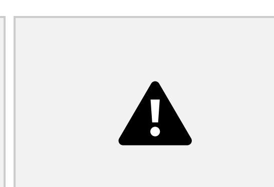


- Machine Translation



- Speech Recognition




RNN - Applications

RNN - Applications

- one-to-many
  - Image Captioning

"black and white dog jumps over bar."

"little girl is eating piece of cake."

Words

[1] Andrej Karpathy, Li Fei-Fei, "Deep Visual-Semantic Alignments for Generating Image Descriptions", CVPR (2015)

e.g. Image Captioning Image to Sequence of

- many-to-one
  - Motion in Video

ords to Sentiment

e.g. Sentiment Classification…
Sequence of Words to Sentiment

CÁM ƠN ĐÃ LẮNG NGHE!