



Deep Learning

Peerapon Vateekul, Ph.D.

peerapon.v@4amconsult.com



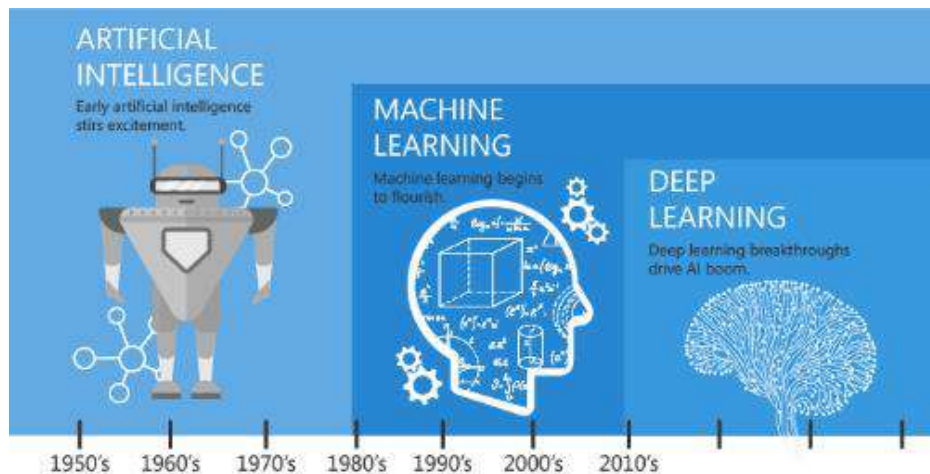
Outlines

- What is Deep Learning (DL)?
- Why now?
- DL application
 - Computer Vision (CV), NLP, Automatic Speech Recognition (ASR)
- DL tools
- Conclusion

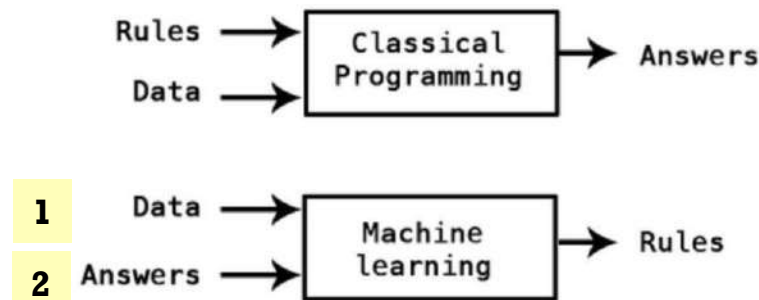


AI = Automation

- 1) Rule-based AI
- 2) Machine Learning (ML)



Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.

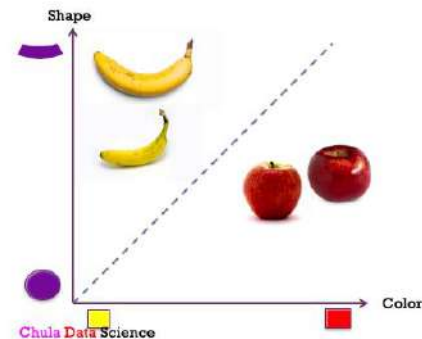


<https://mc.ai/machine-learning-basics-artificial-intelligence-machine-learning-and-deep-learning/>



Handcrafted features

Age	Income	Gender	Province	Corona
25	25,000	Female	Bangkok	Yes
35	50,000	Female	Nontaburi	Yes
32	35,000	Male	Bangkok	No



Can we still tell the features (columns)?





What is Deep Learning (DL)?

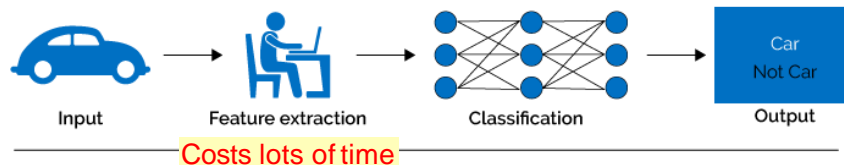


Part of the **machine learning** field of learning representations of data. Exceptional effective at learning patterns.

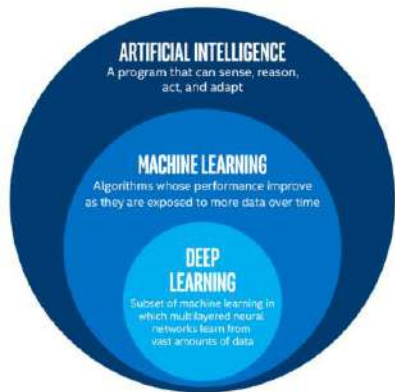
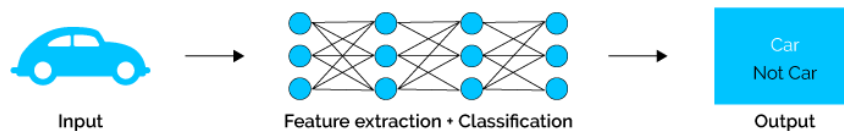


Utilizes learning algorithms that derive meaning out of data by using a **hierarchy** of multiple layers that **mimic the neural networks of our brain**.

Machine Learning



Deep Learning





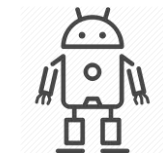
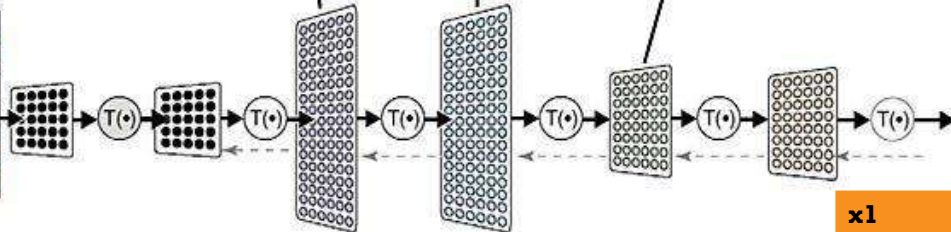
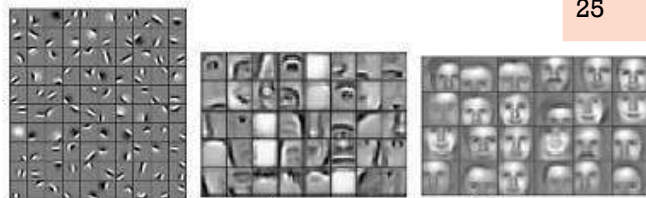
Deep Learning – Basics (cont.)

What did it learn?

A deep neural network consists of a **hierarchy of layers**, whereby each layer **transforms the input data** into more abstract representations (e.g., edge -> nose -> face). The output layer combines those features to make predictions.



Age	Income	Gender	Province	Corona
25	25,000	Female	Bangkok	Yes



x1	x2	x3	x4	Corona
0.7	0.2	-0.5	-0.1	Yes

+ Deep Learning Application



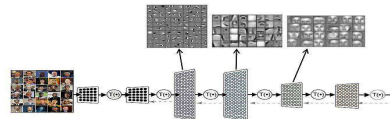
Speech
Recognition



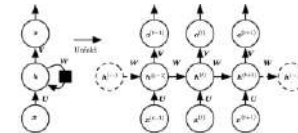
Computer
Vision



Natural Language
Processing



CNN



RNN (LSTM)

IMAGENET

The **ImageNet** project is a large visual database designed for use in visual object recognition software research. Over **14M** URLs of images have been hand-annotated by ImageNet to indicate what objects are pictured on **22K** categories.



ILSVRC

The Image Classification Challenge:

1,000 object classes

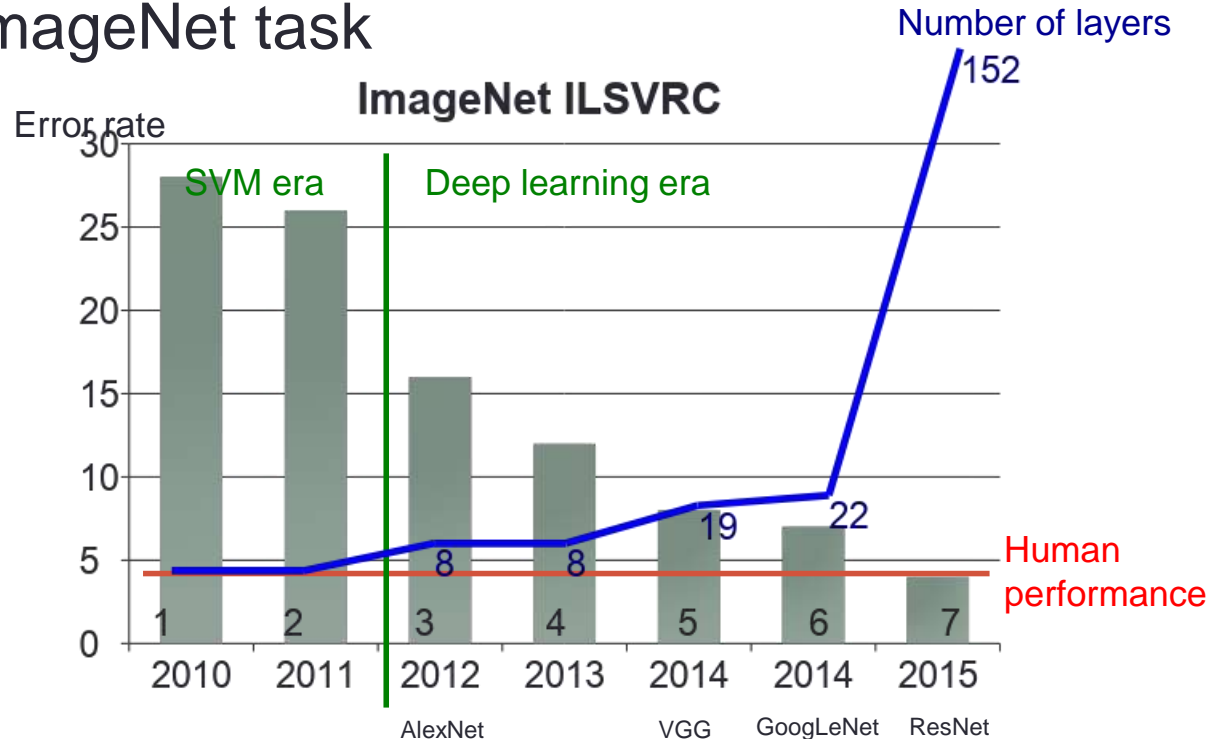
1,431,167 images

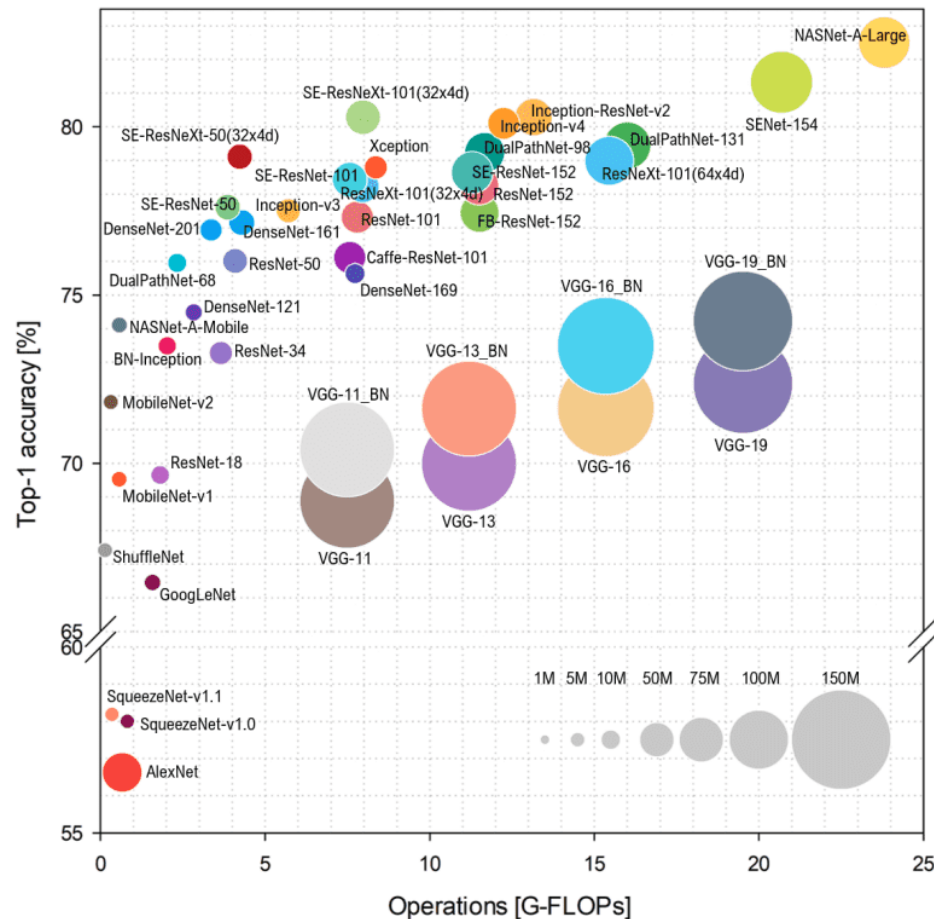
ImageNet 2017 is the last challenge.



Wider and deeper networks (Beyond Human)

- ImageNet task

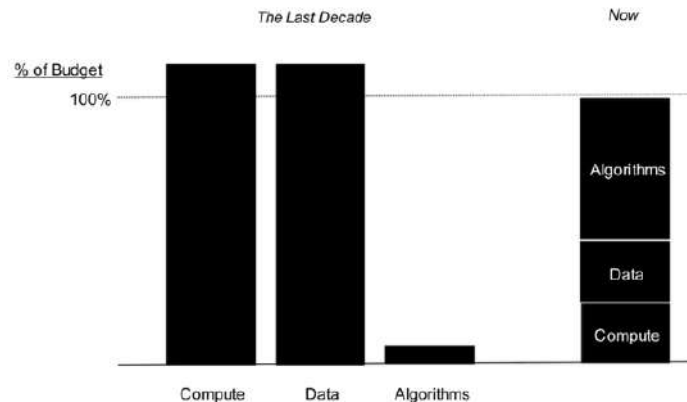




<https://theaisummer.com/cnn-architectures/>

Why now

- Neural Networks has been around since 1990s
- **Big data** – DNN can take advantage of large amounts of data better than other models
- **GPU** – Enable training bigger models possible
- **Deep** – Easier to avoid bad local minima when the model is large





Deep Learning Application



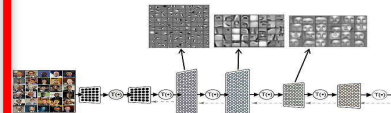
Speech
Recognition



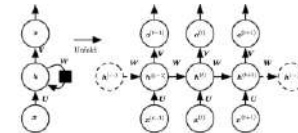
Computer
Vision



Natural Language
Processing



CNN



RNN (LSTM)

Type of image tasks

Semantic Segmentation



GRASS, CAT,
TREE, SKY

No objects, just pixels

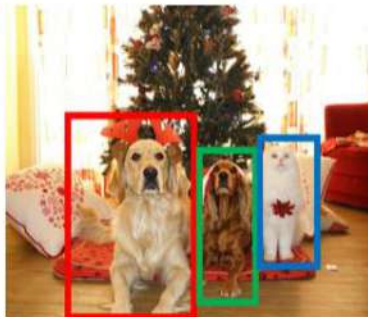
Classification + Localization



CAT

Single Object

Object Detection



DOG, DOG, CAT

Multiple Object

Instance Segmentation



DOG, DOG, CAT

This image is CC0 public domain

PEOPLE COUNTING



<https://www.hikvision.com/th/newsroom/latest-news/2019/people-counting-for-increased-retail-success/>

THE NEW
WORLD



boat 0.40.8

boat 0.55

In Frame: 3
Identified: 3

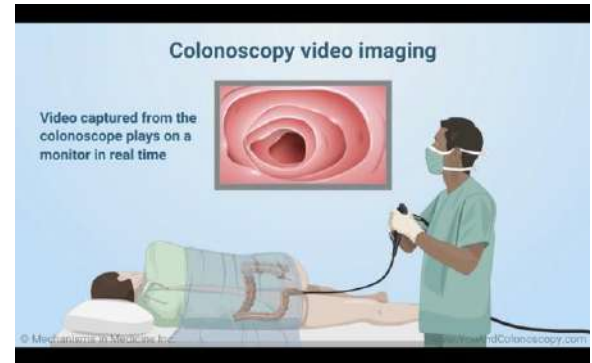
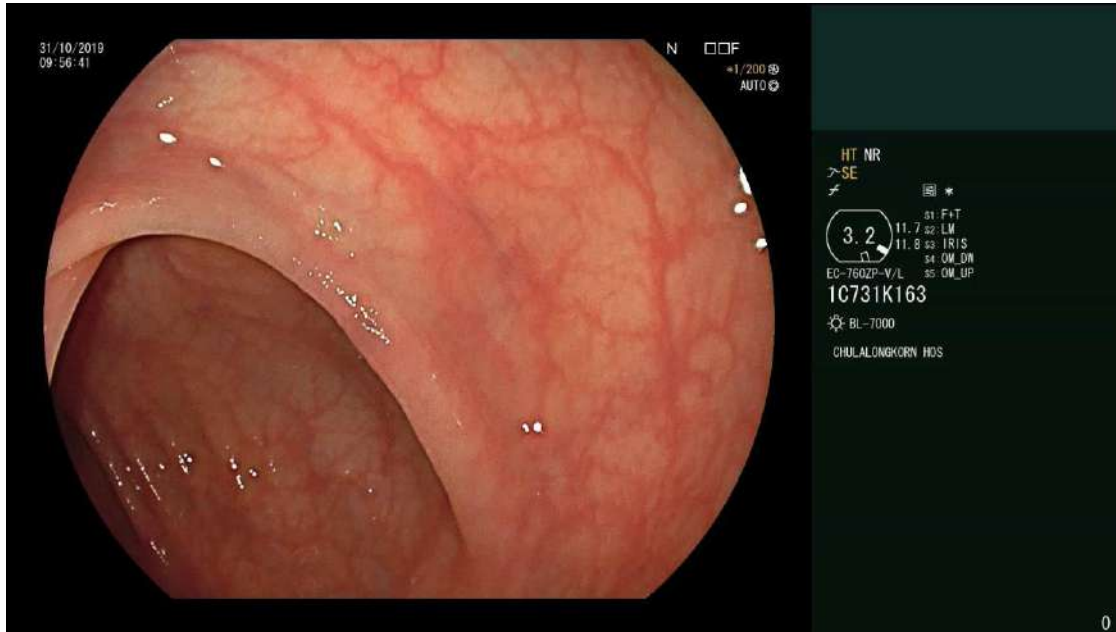
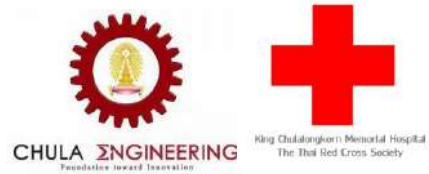


ตัวอย่างความสนุกวันนี้



www.youtube.com/workpointofficial

Smart Medical Diagnosis



car 0.96



traffic cone 0.85



traffic cone 0.89







+ Deep Learning Application



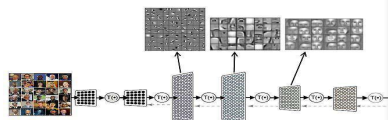
Speech
Recognition



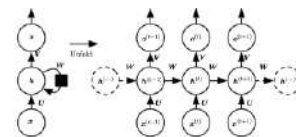
Computer
Vision



Natural Language
Processing



CNN



RNN (LSTM)

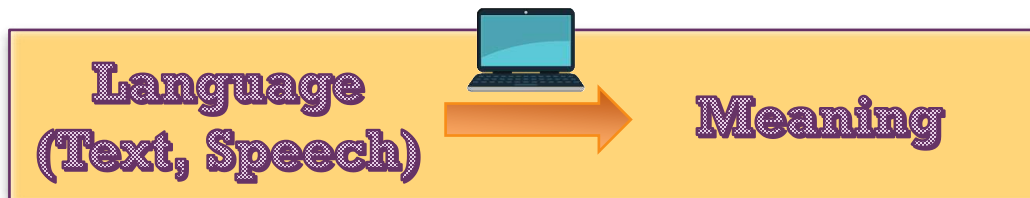


Introduction: Natural Language Processing (NLP)

- Subfield of AI

- GOAL:

- Bridge the gap between **how people communicate** and **what machines understand** in order to perform useful tasks, e.g.
- Making appointments, buying things, question answering, etc.



Goal: intelligent processing of human languages

- **Not just** string matching



Intro (cont.): NLP Pipeline Example (English)

1

Tokenization

- Input: Mr.Smith goes to Washington
- Output: [Mr.Smith, goes, to, Washington]

2

Part of
Speech
tagging

- Input: [Mr.Smith,goes,to,Washington]
- Output:[(Mr.Smith,**NNP**), (goes,**VBZ**), (to,**TO**), (Washington,**NNP**)]

3

NER

- Input:[(Mr.Smith,**NNP**), (goes,**VBZ**), (to,**TO**), (Washington,**NNP**)]
- Output:[(Mr.Smith,**NNP**,**PER**), (goes,**VBZ**,**O**), (to,**TO**,**O**), (Washington,**NNP**,**LOC**)]

4

Application

- e.g., Word Cloud (Named Entity Only), Question Answering (QA)

PENN Part Of Speech Tags

- NNP – proper noun
- VBZ - Verb, 3rd person singular present
- TO – to

Ref:

https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html

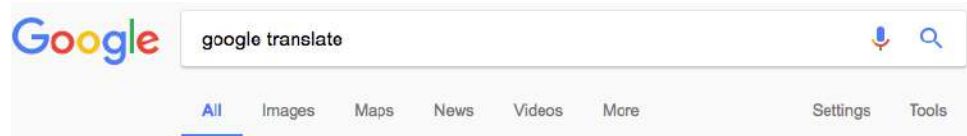
Named Entity Tags

- PER –Person
- LOC – Location
- ORG – Organization
- O – Other





NLP today: Machine Translation (MT)



About 1,180,000,000 results (0.39 seconds)

English ▾

As the new year gets underway, expert commentators give their view on what 2018 holds in store.

Here are three big themes to watch out for over the next 12 months.

Can the stock market rally go on?
The new year has begun with stock markets in the UK and US hitting new record highs.

The Dow Jones Industrial Average rose above 25,000 points for the first time this week, while the broader S&P 500 is also at historic highs.



Thai ▾



เป็นปีใหม่ที่กำลังได้รับการแสดงความคิดเห็นของผู้เชี่ยวชาญให้มุมมองของพวกเขาเกี่ยวกับสิ่งที่ 2018 เก็บในร้าน

ต่อไปนี้เป็นหัวข้อใหญ่สามข้อที่ควรระวังในช่วง 12 เดือนข้างหน้า

การชุมนุมตลาดหุ้นสามารถดำเนินต่อไปได้หรือไม่?
ปีใหม่เริ่มมีตลาดหุ้นในสหราชอาณาจักรและสหรัฐฯ พุ่งสูงเป็นประวัติการณ์

ดัชนีเฉลี่ยอุตสาหกรรมดาวโจนส์ปรับตัวสูงขึ้นกว่า 25,000 จุดเป็นครั้งแรกในสัปดาห์นี้ ขณะที่ดัชนี S & P 500 ที่ใหญ่ขึ้นก็อยู่ในระดับสูงเป็นประวัติการณ์

Markets, Brexit and Bitcoin: 2018's themes

By Chris Johnston
Business reporter

5 January 2018



As the new year gets underway, expert commentators give their view on what 2018 holds in store.

<http://www.bbc.com/news/business-42581934>



NLP today: Question Answering (QA)



IBM Watson wowed the tech industry and a corner of U.S. pop culture with its 2011 win against two of Jeopardy's greatest champions. Here's how IBM pulled it off and a look at what Watson's real career is going to be.

<https://www.techrepublic.com/article/ibm-watson-the-inside-story-of-how-the-jeopardy-winning-supercomputer-was-born-and-what-it-wants-to-do-next/>

Ref: Prof. Regina Barzilay, NLP @MIT

+ Deep Learning Application



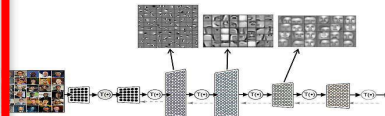
Speech
Recognition



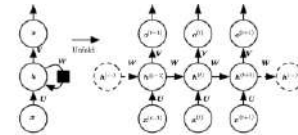
Computer
Vision



Natural Language
Processing



CNN



RNN (LSTM)

Spaces: biodatlab/whisper-thai-demo like 0 Running on T4 Open logs

App Files and versions Community 1 Settings

Linked Models

Transcribe Audio Transcribe YouTube

Whisper Demo Thai: Transcribe Audio

Transcribe long-form microphone or audio inputs with the click of a button! Demo uses the the fine-tuned checkpoint [biodatlab/whisper-medium-th-cmv11](#) and 🤗 Transformers to transcribe audio files of arbitrary length.

🔊 microphone

✎ ✕

▶ 0:00 🔊

📎 file_upload

Drop Audio Here

- or -

Click to Upload

Clear

Submit

queue: 2/2 | 147.2/416.5





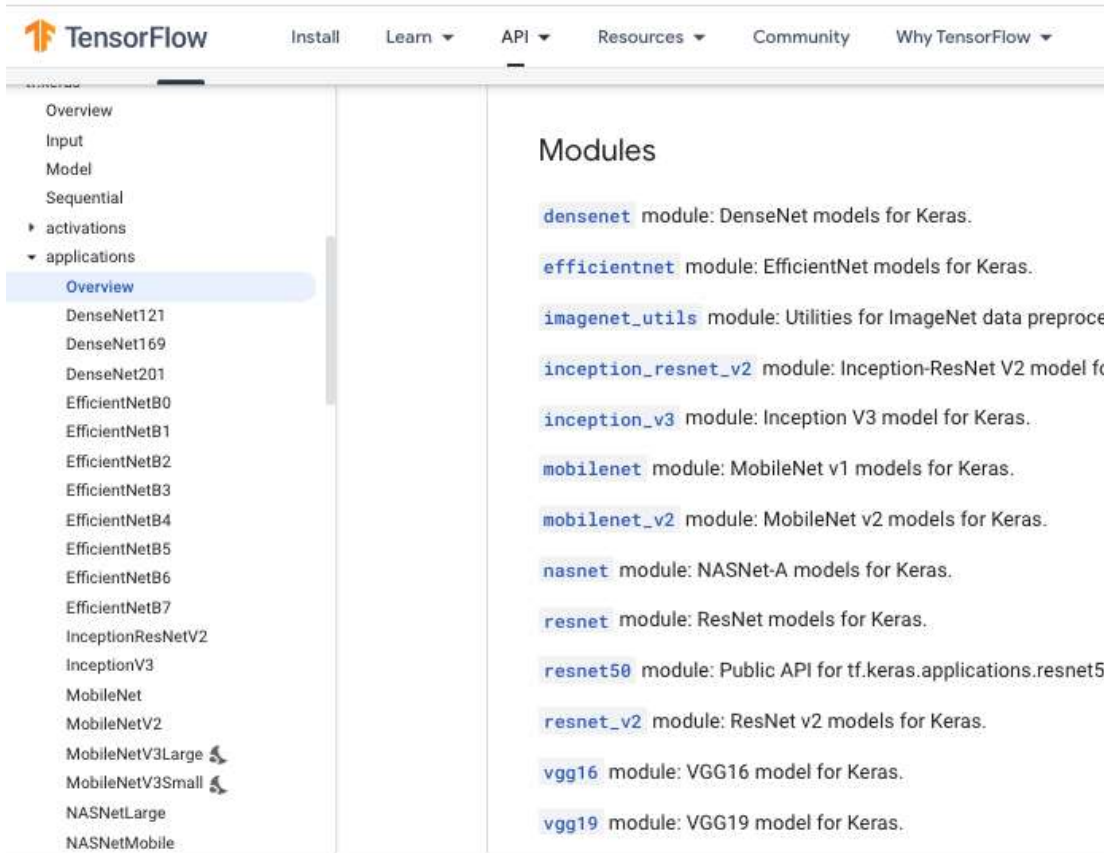
Deep Learning tools



+



Tools for Computer Vision



The screenshot shows the TensorFlow website. The top navigation bar includes links for 'Install', 'Learn', 'API', 'Resources', 'Community', and 'Why TensorFlow'. The left sidebar contains a list of categories: 'Overview', 'Input', 'Model', 'Sequential', 'activations', and 'applications'. The 'applications' category is expanded, showing a list of models including DenseNet, EfficientNet, Inception, MobileNet, and NASNet. The main content area is titled 'Modules' and lists various Keras modules with their descriptions.

TensorFlow

Install Learn API Resources Community Why TensorFlow

Overview

Input

Model

Sequential

activations

applications

Overview

DenseNet121

DenseNet169

DenseNet201

EfficientNetB0

EfficientNetB1

EfficientNetB2

EfficientNetB3

EfficientNetB4

EfficientNetB5

EfficientNetB6

EfficientNetB7

InceptionResNetV2

InceptionV3

MobileNet

MobileNetV2

MobileNetV3Large

MobileNetV3Small

NASNetLarge

NASNetMobile

Modules

densenet module: DenseNet models for Keras.

efficientnet module: EfficientNet models for Keras.

imagenet_utils module: Utilities for ImageNet data preprocessing & prediction decoding.

inception_resnet_v2 module: Inception-ResNet V2 model for Keras.

inception_v3 module: Inception V3 model for Keras.

mobilenet module: MobileNet v1 models for Keras.

mobilenet_v2 module: MobileNet v2 models for Keras.

nasnet module: NASNet-A models for Keras.

resnet module: ResNet models for Keras.

resnet50 module: Public API for tf.keras.applications.resnet50 namespace.

resnet_v2 module: ResNet v2 models for Keras.

vgg16 module: VGG16 model for Keras.

vgg19 module: VGG19 model for Keras.

Model 1

- VGG19 (random initialized weights) + 2 Dense layers + Output layer

```
1 base_model = VGG19(weights=None, include_top=False, input_shape=(224, 224, 3))
2
3 for layer in base_model.layers:
4     layer.trainable = True
5
6 x = base_model.output
7 x = Flatten()(x)
8 x = Dense(1024)(x)
9 x = Dropout(0.5)(x)
10 x = Dense(512)(x)
11 x = Dropout(0.5)(x)
12 output = Dense(num_class, activation='softmax')(x)
13
```



Tools for Thai NLP



PyThaiNLP

python 3.6 pypi v2.1.2 downloads/month 15k License Apache 2.0 license scan passing build passing build passing
code quality A coverage 91% Launch Quick Start Guide on Google Colab DOI 10.5281/zenodo.3595968

Thai Natural Language Processing in Python.

PyThaiNLP is a Python package for text processing and linguistic analysis, similar to `nlTK` but with focus on Thai language.

News

We are conducting a 2-minute survey to know more about your experience using the library and your expectations regarding what the library should be able to do. Take part in this survey: <https://forms.gle/aLdSHnvkNuK5CFy9>

This is a document for development branch (post 2.1). Things will break.

WangchanBERTa โมเดลประมวลผลภาษาไทยที่ใหญ่และก้าวหน้าที่สุดในขณะนี้



VISTEC-depa AI Research Institute of Thailand

Follow

Jan 24 · 5 min read



เปิดให้ทุกคนใช้ฟรีโดย AIResearch.in.th และ VISTEC ภายใต้สัญญาอนุญาต CC-BY-SA 4.0





ปล่อยโมเดล

Thonburian Whisper

Automatic Speech Recognition in Thai

ถอดความจากเสียงพูดภาษาไทยที่พัฒนาต่อจากโมเดลของ OpenAI

Conclusion

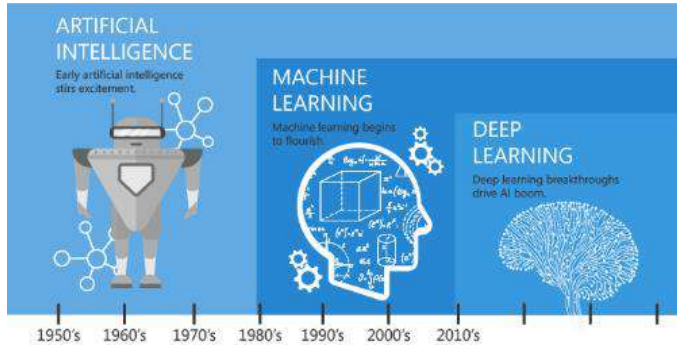
MACHINE LEARNING
DRIVING PATTERN

DISTANCE

ACCELERATION

REACTION SPEED

Conclusion



Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.



Speech Recognition

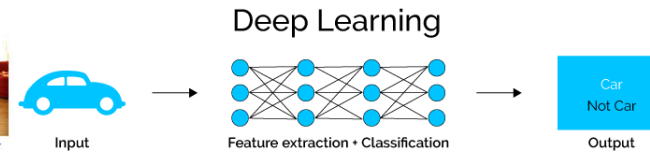
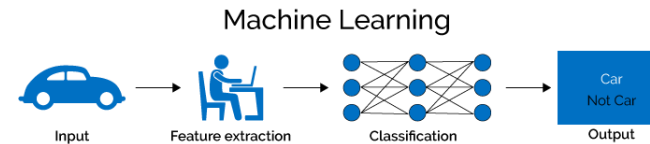
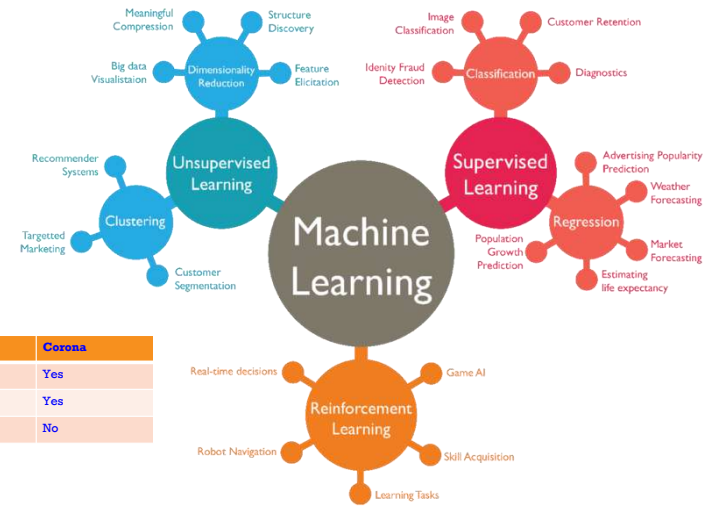
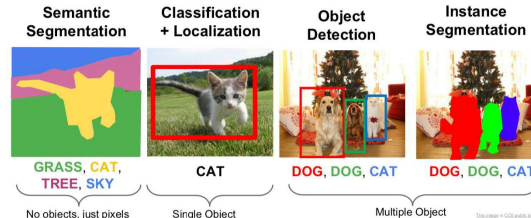


Computer Vision



Natural Language Processing

Age	Income	Gender	Province	Corona
25	25,000	Female	Bangkok	Yes
35	50,000	Female	Nontaburi	Yes
32	35,000	Male	Bangkok	No



A top-down view of a white spiral-bound notebook resting on a light-colored wooden surface. The notebook is open to a page with the words 'THANK YOU FOR YOUR ATTENTION' written in large, bold, black, hand-drawn capital letters. To the top left of the notebook, a portion of a green leafy plant is visible. To the top right, a pair of black-rimmed glasses lies on the surface. To the bottom right of the notebook, a silver and gold ballpoint pen is positioned diagonally.

**THANK YOU
FOR YOUR
ATTENTION**