

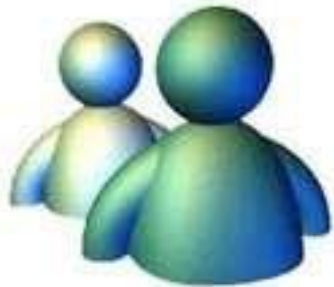


Deep Learning for Human Language Processing

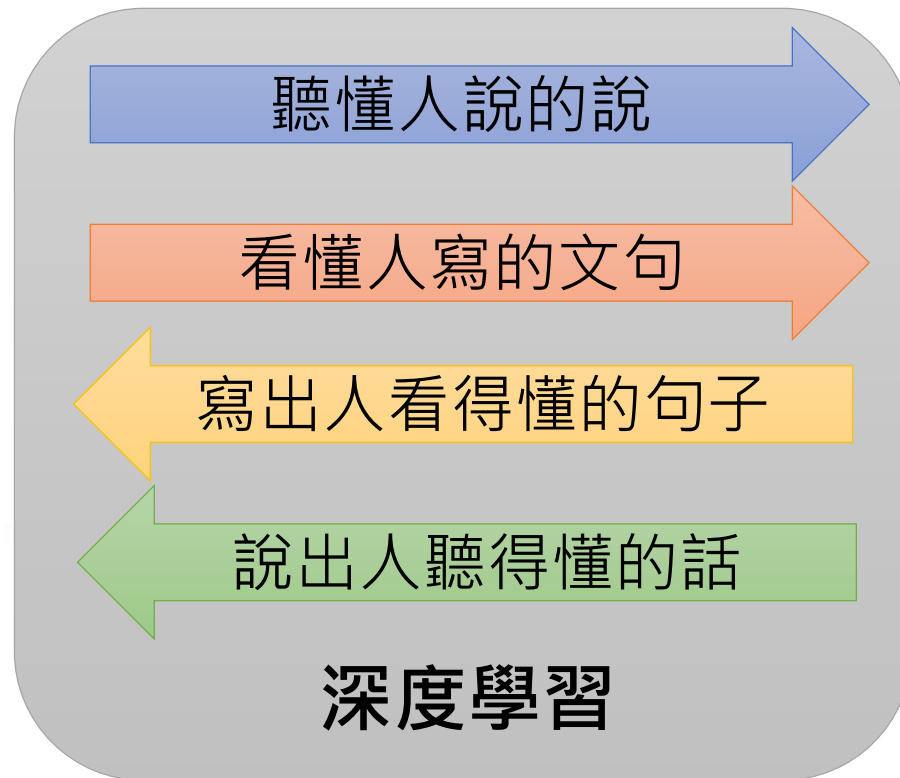
HUNG-YILEE
李宏毅

What is this course about?

- 深度學習與人類語言處理 (Deep Learning for Human Language Processing)



Human



Computer

What is this course about?

- 深度學習與人類語言處理 (Deep Learning for Human Language Processing)
- 自然語言處理 (Natural Language Processing, NLP)
 - A language that has developed naturally in use (e.g. Chinese, English)
 - As contrasted with an artificial language (e.g. JAVA, Python)
 - Natural Language can be Speech or Text

這門課也可以叫「深度學習與自然語言處理」

Why not???

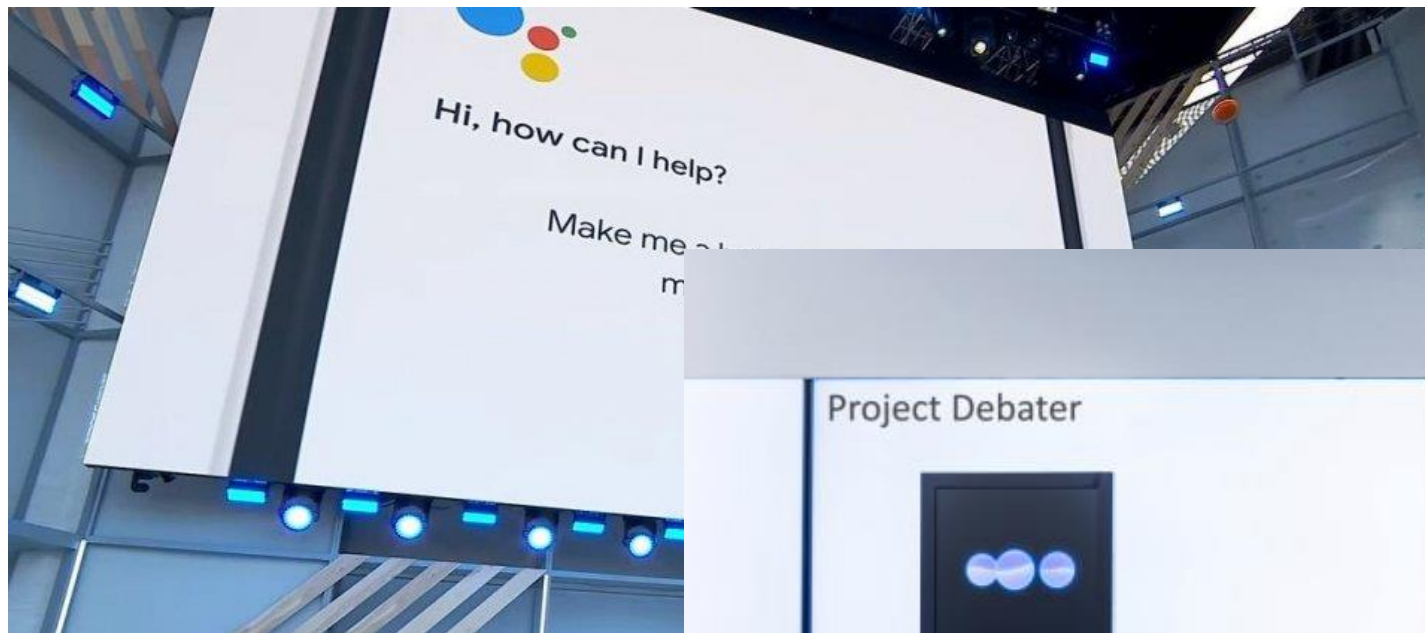
What is this course about?

- In this course, Text v.s. Speech = 5 : 5
- Most NLP textbook and course mainly focus on text (Text v.s. Speech = 9 : 1)

所以這門課叫做「深度學習與人類語言處理」

- Speech processing is NOT only speech recognition.
- Only 56% languages have written form (Ethnologue, 21st edition)
 - We don't always know if the existing writing systems are widely used.

Human Language Processing is popular!



Google Duplex (2018)



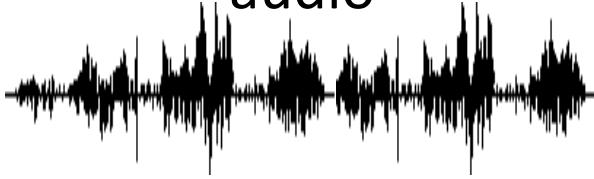
IBM Project Debater (2019)

Ref: <https://thejohnfox.com/long-sentences/>

Ref: https://en.wikipedia.org/wiki/Longest_English_sentence

Human Language is complex

audio



1 second has 16K sample points

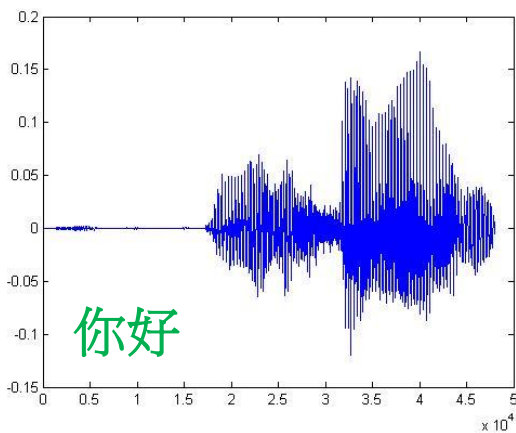
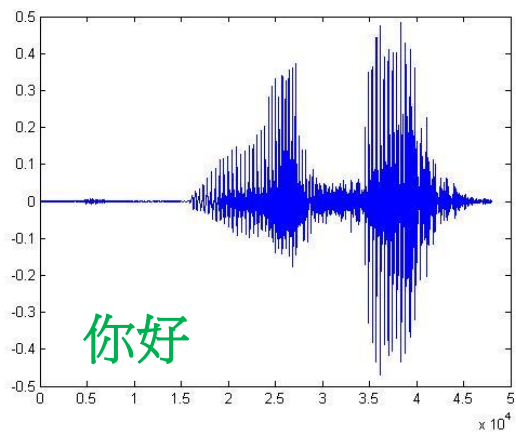
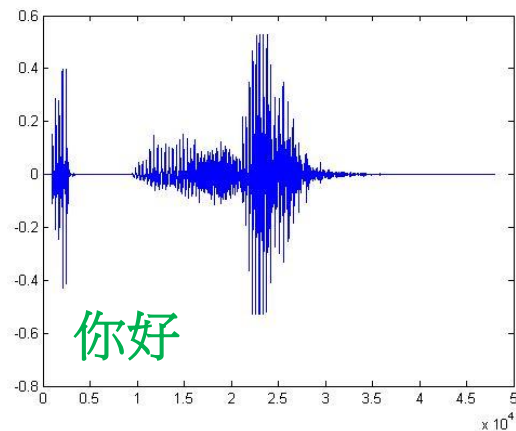
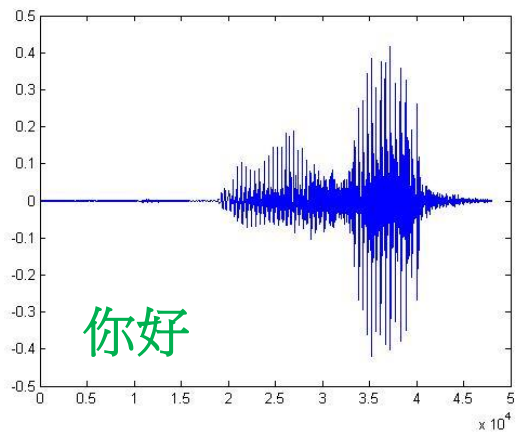
Each point has 256 possible values.

NO MAN EVER STEPS
IN THE SAME RIVER
TWICE FOR IT'S NOT
THE SAME RIVER AND
HE'S NOT THE SAME

M A N

古希臘哲學家赫拉克利特(Heraclitus)

HERACLITUS



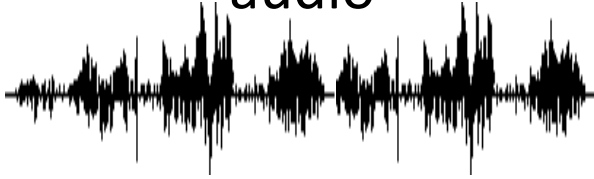
也沒有人可以說同一段話兩次

Ref: <https://thejohnfox.com/long-sentences/>

Ref: https://en.wikipedia.org/wiki/Longest_English_sentence

Human Language is complex

audio



1 second has 16K sample points

Each point has 256 possible values.

text



William Faulkner, “Absalom, Absalom.”:
“Just exactly like Father if Father had
known” (1289 words)

Jonathan Coe's *The Rotters' Club* has a
sentence with 13,955 words (2014)

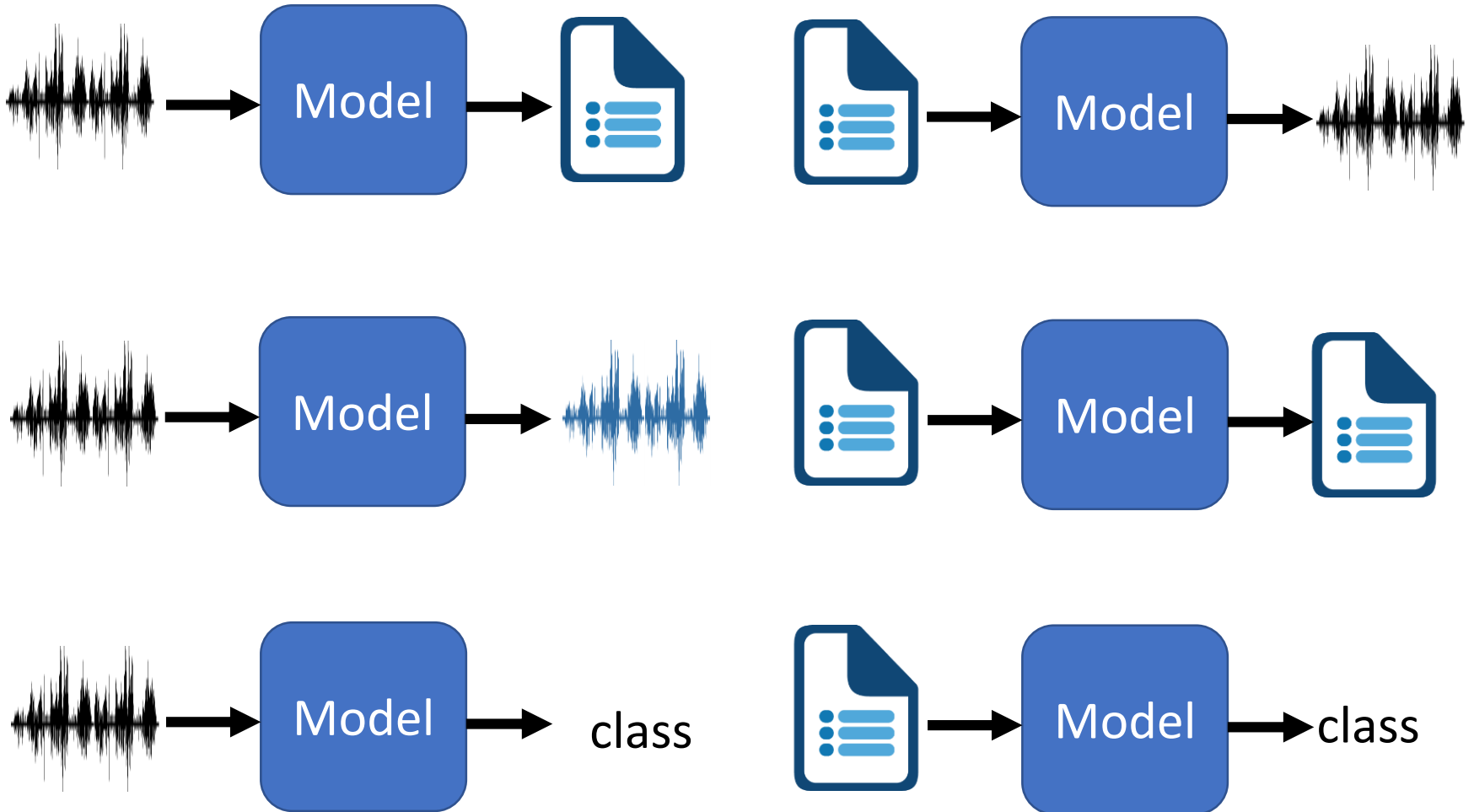
Faulkner wrote, “Just exactly like Father ...”

Pinker said Faulkner wrote, “Just exactly like Father ...”

Who cares that Pinker said Faulkner wrote, “Just exactly like Father ...”

*The Language Instinct: How the Mind
Creates Language* (Steven Arthur Pinker)

One slide for this course



硬 train 一發的故事: <https://youtu.be/F1vek6ULo9w>

What is the model?

Model = Deep Network

遇到問題用 deep learning 「硬 train 一發」就對了

沒有「硬 train 一發」
無法解決的問題

如果有 ...
那只是你訓練資料和
GPU 不夠多而已



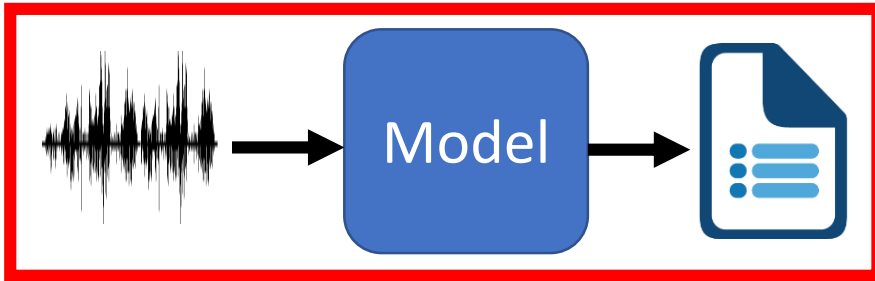


THE DAY AFTER TOMORROW
WHERE WILL YOU BE?

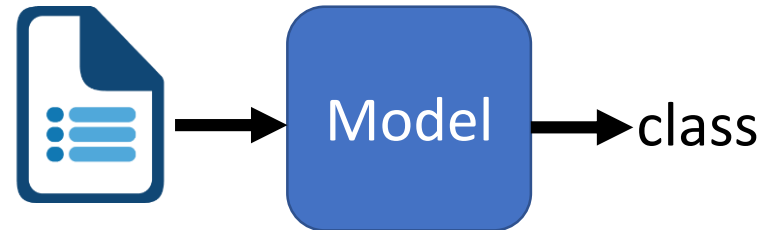
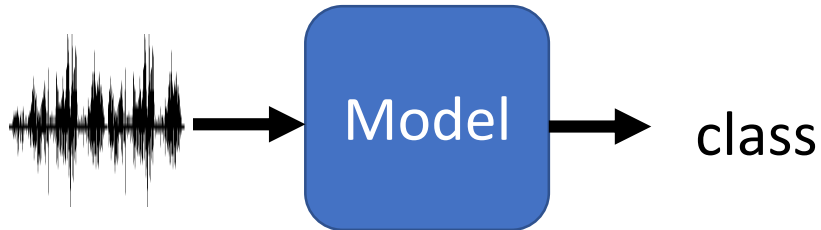
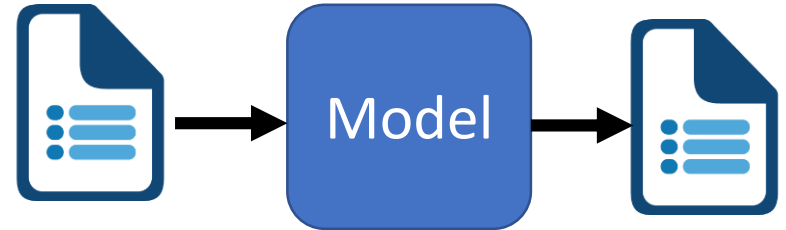
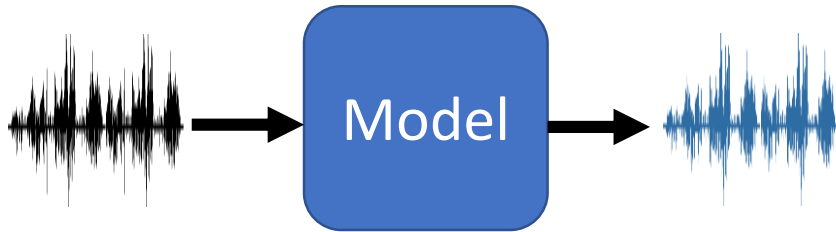
IN THEATERS WORLDWIDE MAY 28, 2004

「硬 train 一發」過後
人類語言處理
的下一步

One slide for this course



Speech Recognition



Automatic Speech Recognition (ASR)

<https://ai.googleblog.com/2019/03/-all-neural-on-device-speech.html>

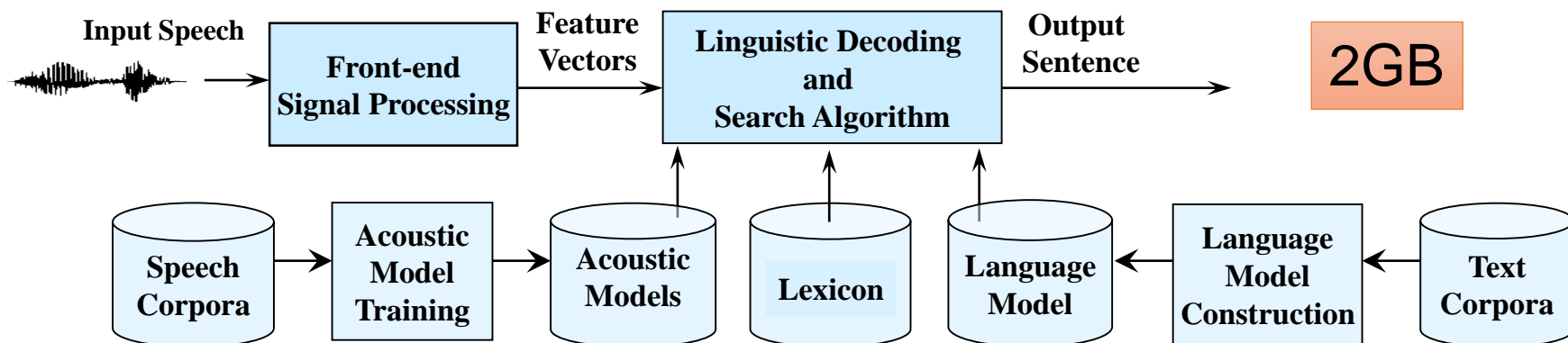
It is not the seq2seq you know!

End-to-end

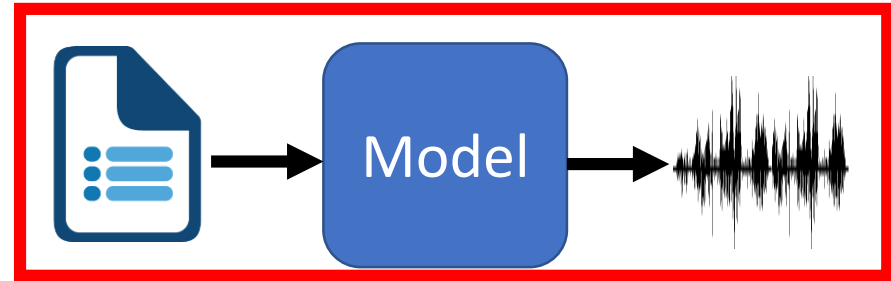
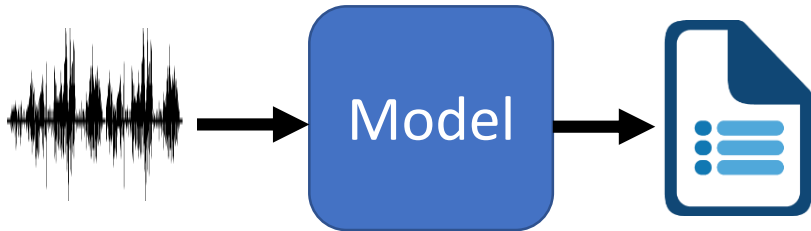
80MB



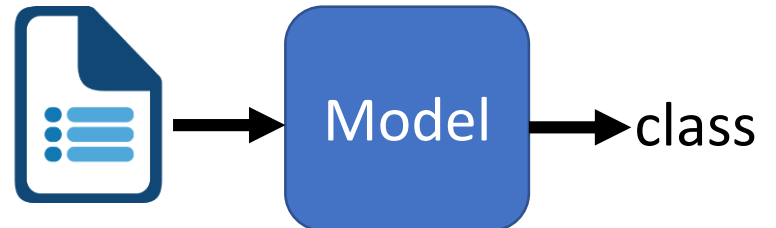
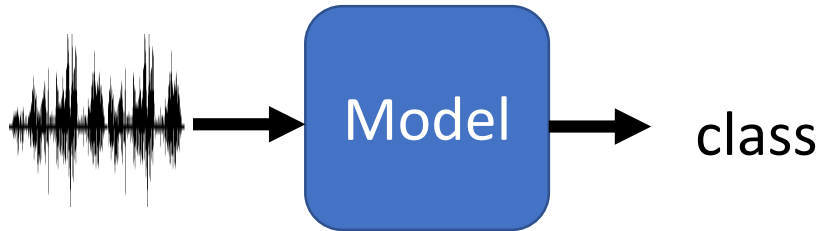
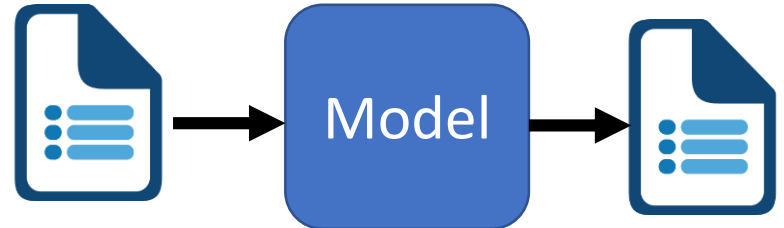
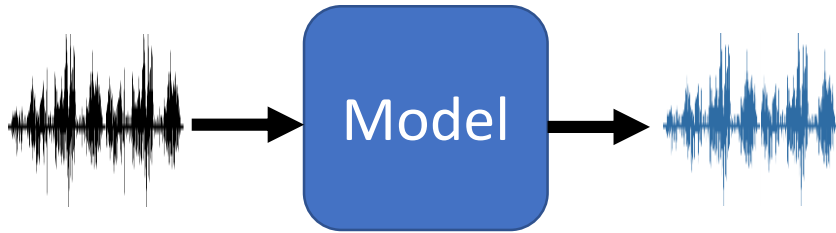
Traditional Speech Recognition



One slide for this course



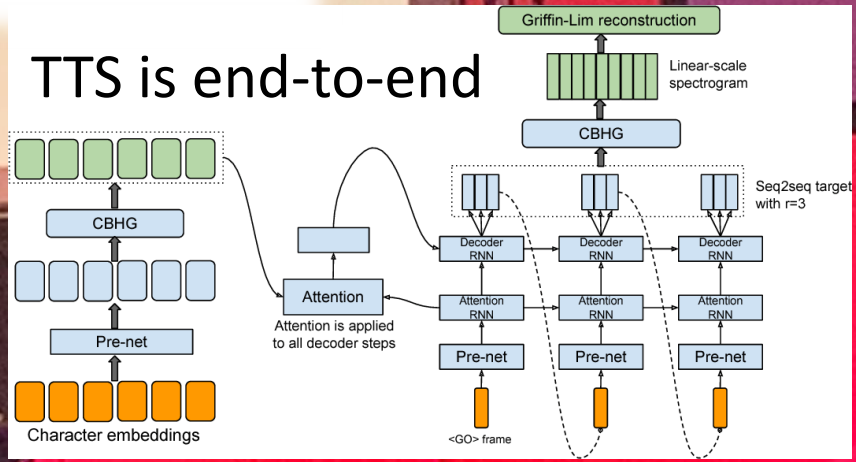
Text-to-Speech Synthesis



Famous words in speech technology (1980s)

“Every time I fire a linguist,
the performance of the speech recognizer goes up”
by Frederick Jelinek

(Keiichi Tokuda, keynote,
INTERSPEECH'19)



All the problems solved?



高雄發大財我現在要出征



發財發財發財發財



發財發財發財

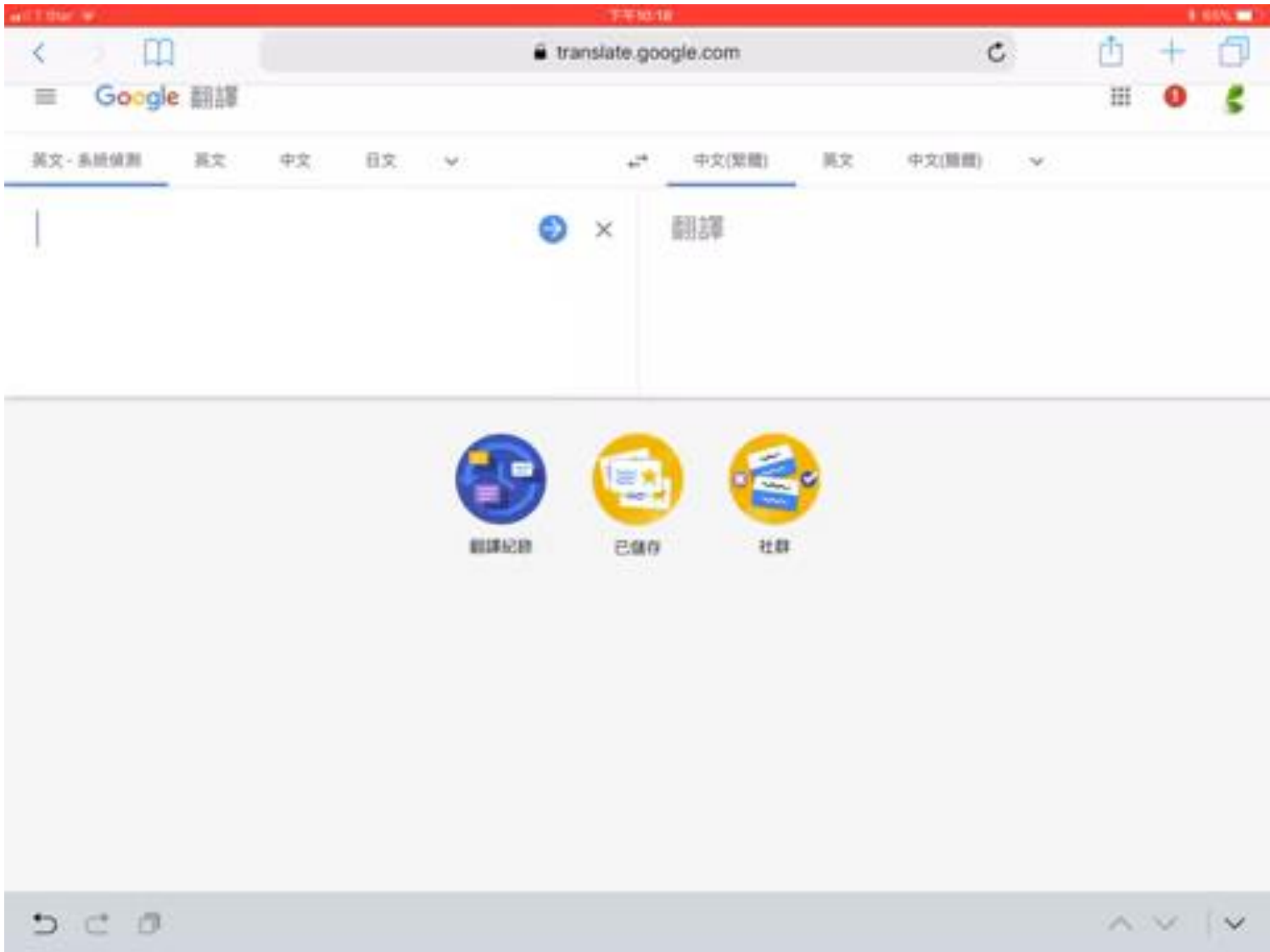


發財發財



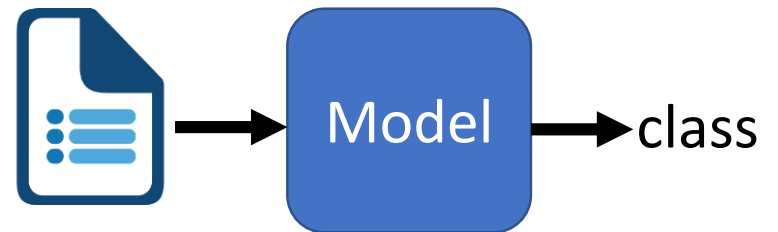
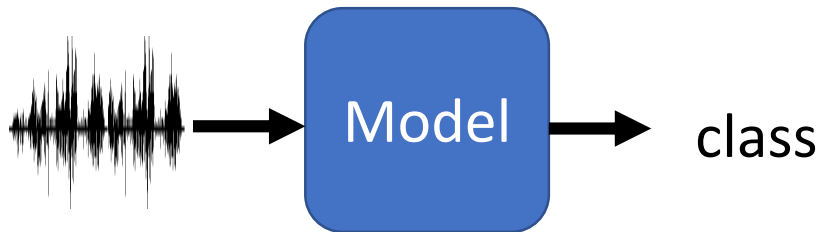
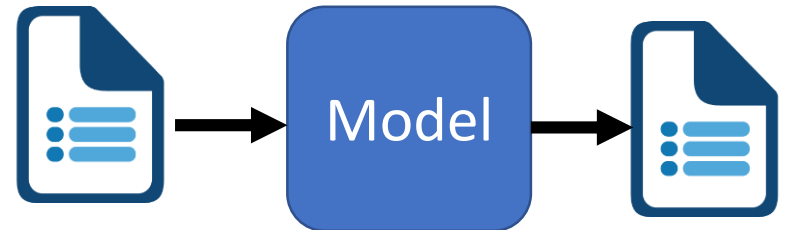
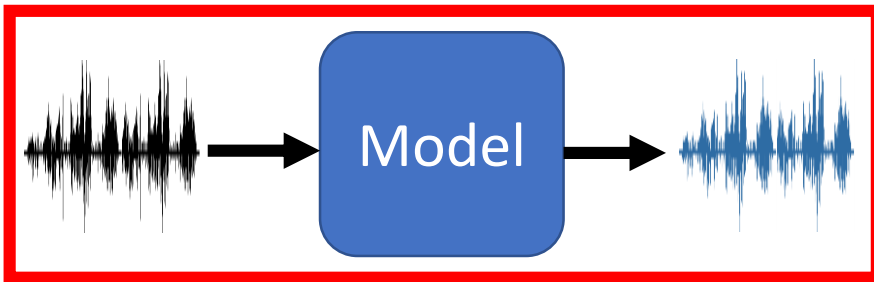
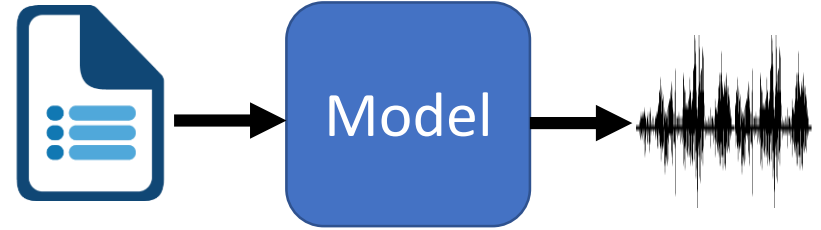
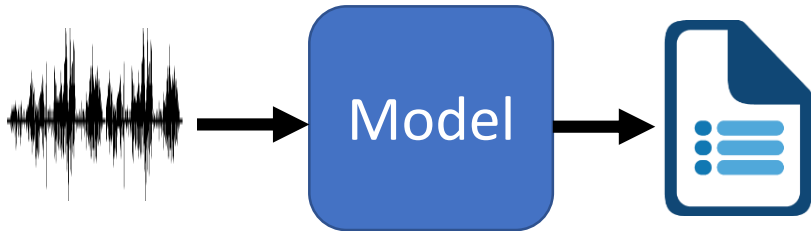
發財

It has happened in real applications



蘋果仁頻道：<https://www.youtube.com/watch?v=EwbTlnUkctM>
This problem is found in 2018.02. It is already fixed.

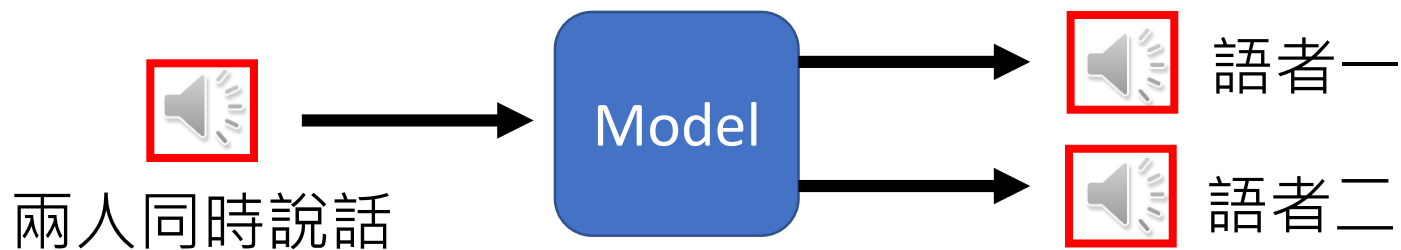
One slide for this course



Speech Separation

感謝 孫凡耕同學、施順耀
同學提供實驗結果

- 雞尾酒會效應 (cocktail party effect)

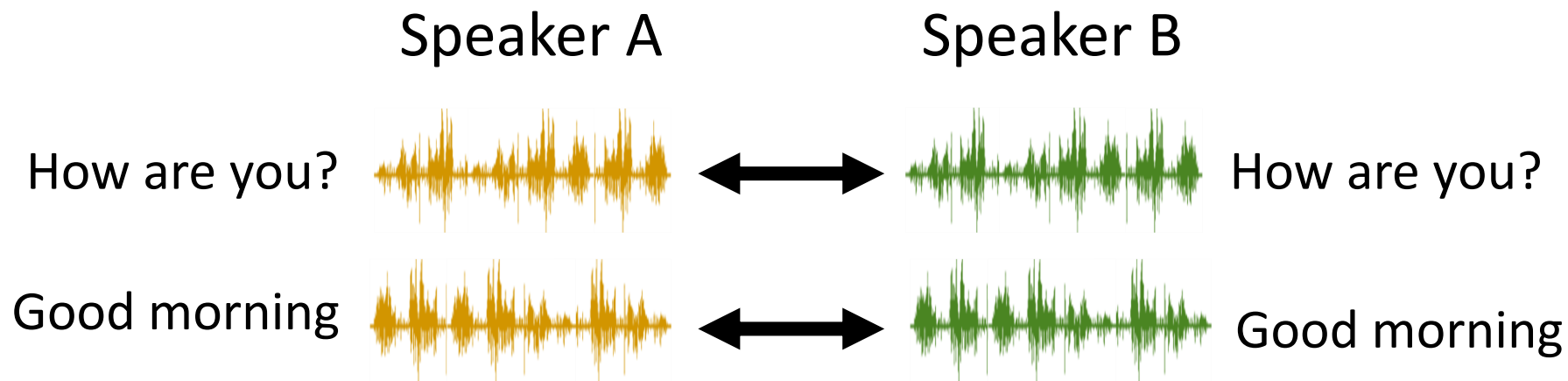


上面結果連 Fourier Transform 都沒有用上
只有用深度學習 “硬train一發”

Voice Conversion



要硬 train 一發你需要



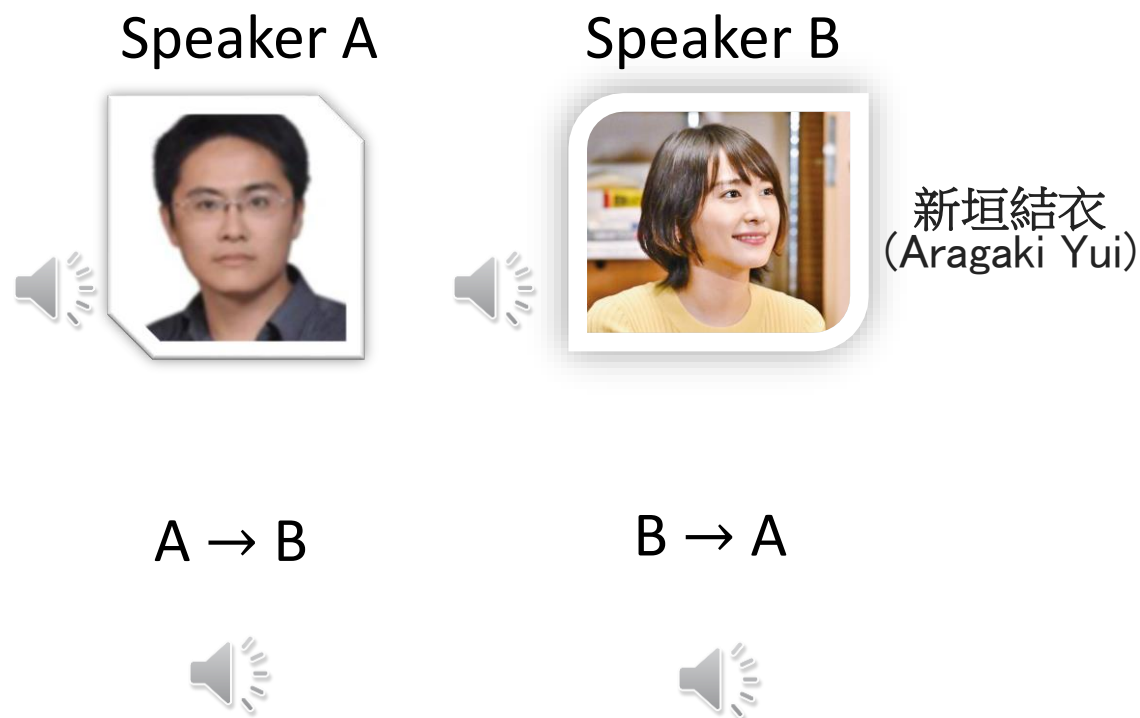
能不能



Speakers A and B are talking about completely different things.

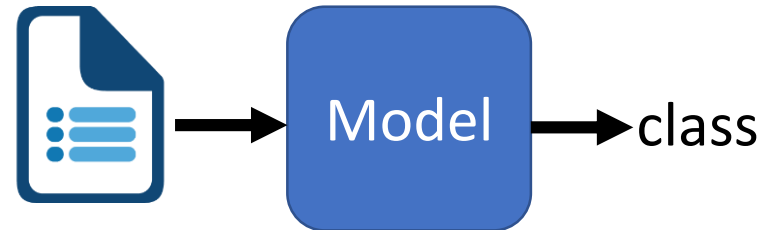
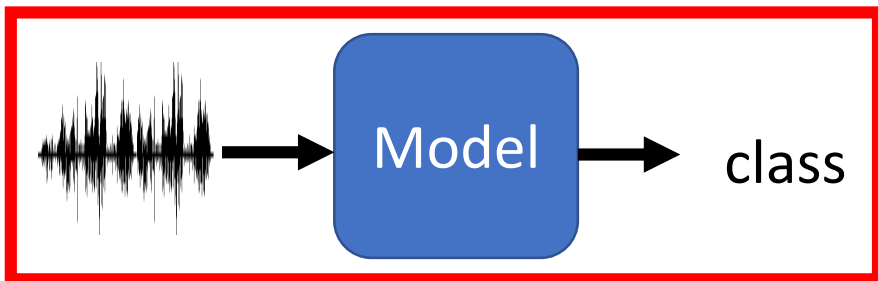
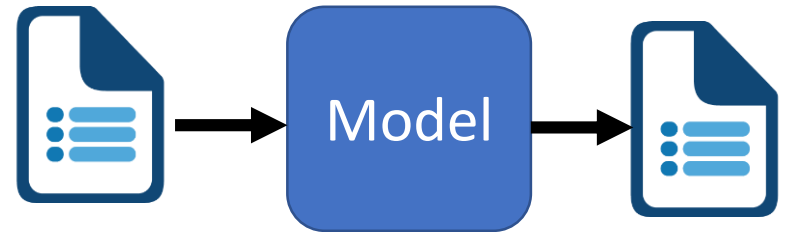
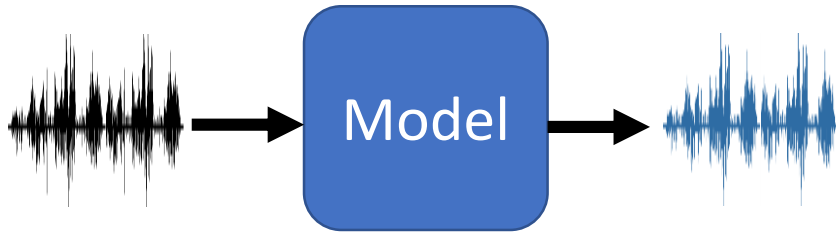
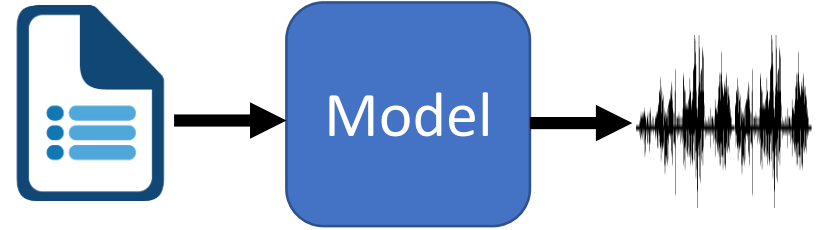
Unsupervised Voice Conversion

- Only one utterance from each speaker (**one-shot learning**)



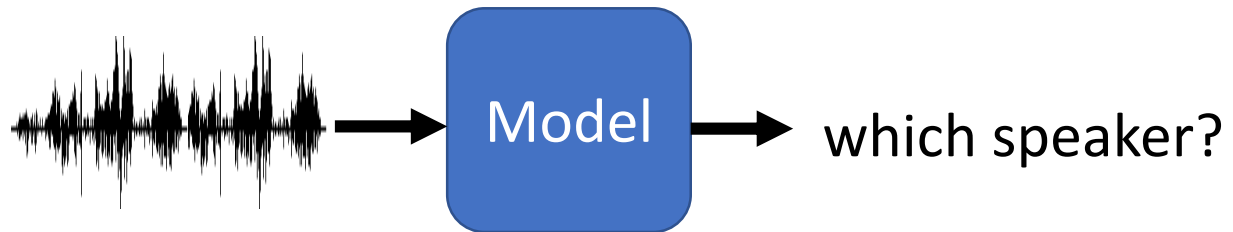
感謝解正平同學
提供實驗結果

One slide for this course



Input Audio, Output Class

Speaker Recognition



"Alexa"

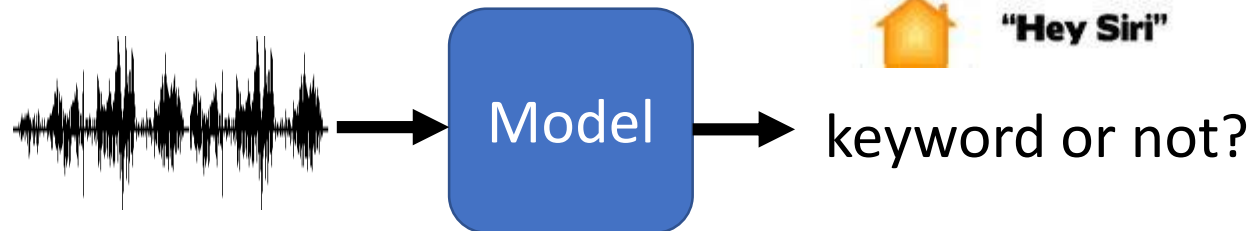


"Ok Google"



"Hey Siri"

Keyword Spotting



Wake up words

- 2017.01, in Dallas, Texas
- A six-year-old asked her Amazon Echo “can you play dollhouse with me and get me a dollhouse?”
- The device orders a KidKraft Sparkle mansion dollhouse.
- TV station CW-6 in San Diego, California, was doing a morning news segment
 - Anchor Jim Patton said, “I love the little girl saying, ‘Alexa ordered me a dollhouse.’ ”

<https://www.foxnews.com/tech/6-year-old-accidentally-orders-high-end-treats-with-amazons-alexa>

<https://www.theverge.com/2017/1/7/14200210/amazon-alexa-tech-news-anchor-order-dollhouse>

Wake up words

2017.04



Whopper

From Wikipedia, the free encyclopedia

This is an **old revision** of this page, as edited by **Julietdeltalima** ([talk](#) | [contribs](#)) at 17:50, 4 April 2017 (*Reverted to revision 7738099 WP:NPOV changes from encyclopedic language to marketingese. (TW)*). The present address (URL) is a [permanent link](#) to this revision. **current revision**.

[\(diff\)](#) ← [Previous revision](#) | [Latest revision \(diff\)](#) | [Newer revision](#) → [\(diff\)](#)

This article is about the hamburger. For the candy, see [Whoppers](#). For other uses, see [Whopper \(disambiguation\)](#).

The **Whopper** is the signature hamburger product sold by the international [fast-food restaurant](#) chain [Burger King](#) and its Australian franchise [Hungry Jack's](#). Introduced in 1957, it has undergone several reformulations including resizing and

Whopper

From Wikipedia, the free encyclopedia

This is an **old revision** of this page, as edited by **Fermachado123** ([talk](#) | [contribs](#)) at 18:14, 4 April 2017 (*updated information on the de nowadays.*). The present address (URL) is a [permanent link](#) to this revision, which may differ significantly from the **current revision**.

[\(diff\)](#) ← [Previous revision](#) | [Latest revision \(diff\)](#) | [Newer revision](#) → [\(diff\)](#)

This article is about the hamburger. For the candy, see [Whoppers](#). For other uses, see [Whopper \(disambiguation\)](#).

The Whopper is a burger, consisting of a flame-grilled patty made with 100% beef with no preservatives, no fillers and is topped with daily sliced tomatoes and onions, fresh lettuce, pickles, ketchup and mayo, served on a soft sesame seed bun. It is the signature hamburger product sold by the international [fast-food restaurant](#) chain [Burger King](#) and its

Fermachado123 is the username of Burger King's marketing chief

Whopper

From Wikipedia, the free encyclopedia

This is an old revision of this page, as edited by 74.108.27.250 (talk) at 15:10, 11 April 2017. The present address (URL) is a permanent link which may differ significantly from the current revision.

(diff) ← Previous revision | Latest revision (diff) | Newer revision → (diff)

This article is about the hamburger. For the candy, see [Whoppers](#). For other uses, see [Whopper \(disambiguation\)](#).

The Whopper is a burger, consisting of a flame-grilled patty made with 100% medium-sized child with no preservatives or fillers, topped with sliced tomatoes, onions, lettuce, pickles, ketchup, and mayonnaise, served on a sesame-seed bun.

Whopper

From Wikipedia, the free encyclopedia

This is an old revision of this page, as edited by 2600:387:5:803::88 (talk) at 16:35, 12 April 2017 (*Fixed typo*). The present address (URL) is a permanent link which may differ significantly from the current revision.

(diff) ← Previous revision | Latest revision (diff) | Newer revision → (diff)

This article is about the hamburger. For the candy, see [Whoppers](#). For other uses, see [Whopper \(disambiguation\)](#).

The Whopper is a burger, consisting of a flame-grilled patty made with 100% rat and toenail clippings with no preservatives or fillers, topped with sliced tomatoes, onions, lettuce, pickles, ketchup, and mayonnaise, served on a sesame-seed bun.

Whopper

From Wikipedia, the free encyclopedia

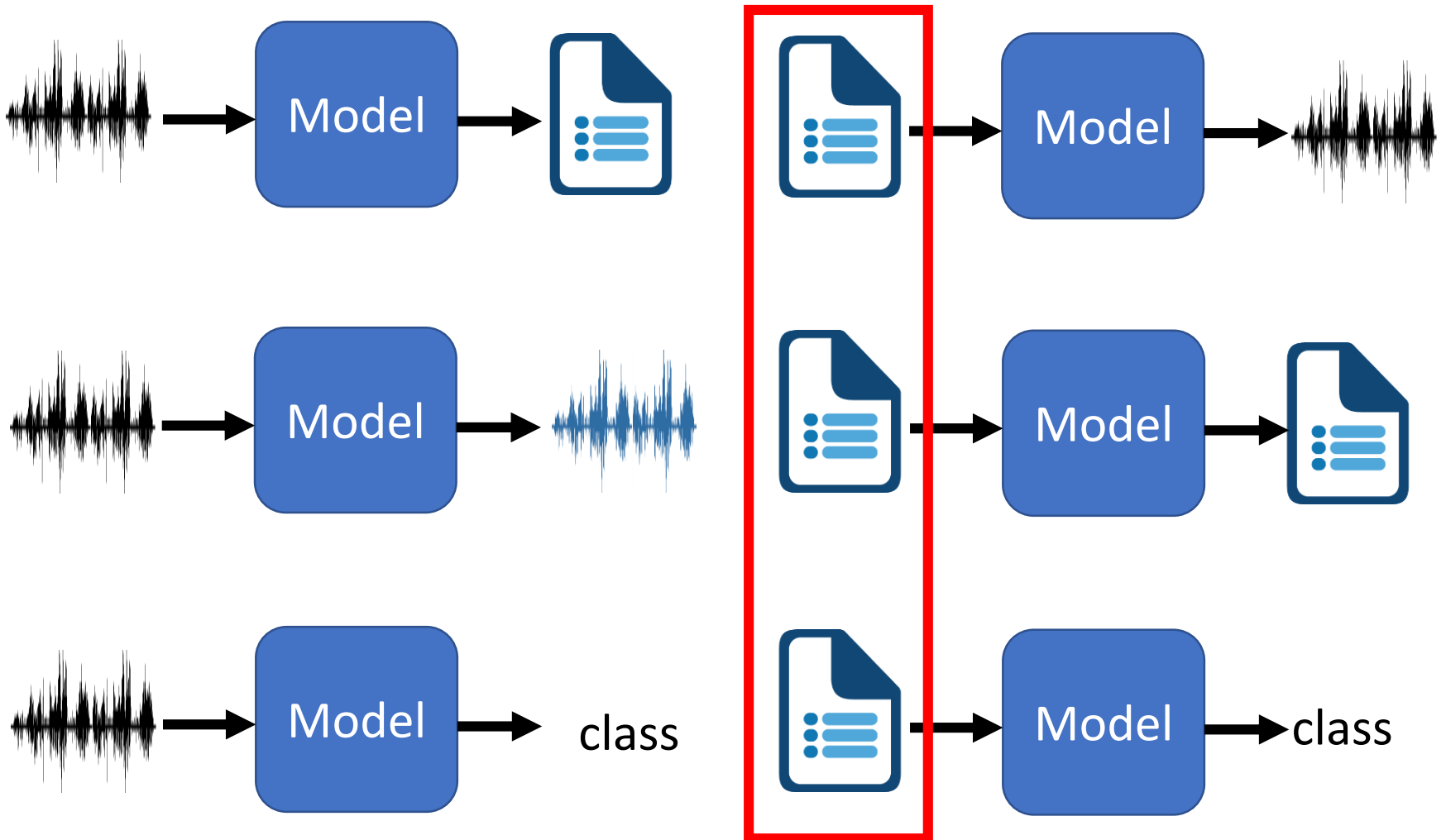
This is an old revision of this page, as edited by [185.58.25.215 \(talk\)](#) at 16:26, 12 April 2017 (*erm*). The present address (URL) is a page that differs significantly from the **current revision.**

[\(diff\)](#) ← [Previous revision](#) | [Latest revision \(diff\)](#) | [Newer revision](#) → [\(diff\)](#)

This article is about the hamburger. For the candy, see [Whoppers](#). For other uses, see [Whopper \(disambiguation\)](#).

The **Whopper** is a signature hamburger product sold by the international fast-food restaurant chain [Burger King](#) and its Australian franchise [Hungry Jack's](#). Introduced in 1957^[*citation needed*], it has undergone several reformulations including resizing and bread changes, yet it remains far inferior to the Big Mac. The burger is one of the best known products in the fast food industry; it is so well known that Burger King bills itself as *the Home of the Whopper* in its [advertising](#) and signage. Additionally, the company uses the name in its high-end concept, the [BK Whopper Bar](#). Due to its place in the marketplace, the Whopper has prompted Burger King's competitors, mainly [McDonald's](#) and [Wendy's](#), to try to develop similar products designed to compete with it.

One slide for this course





是

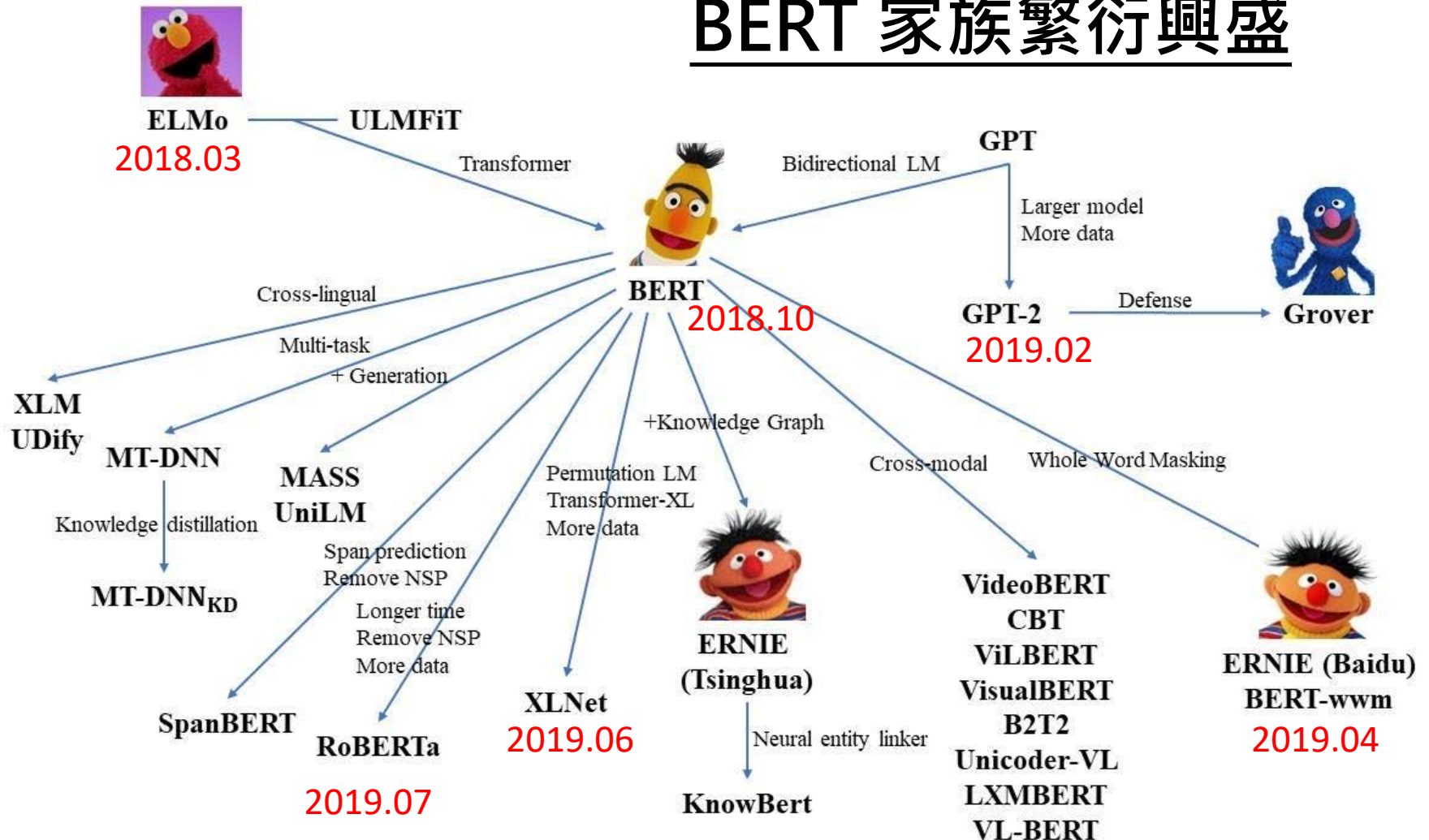
BERT
跟他的好朋友們

瑪利亞之牆

= 之前 NLP 的 state-of-the-art

進展超乎想像 ...

BERT 家族繁衍興盛



The models are become larger and larger ...



BERT
(340M)

ELMO
(94M)

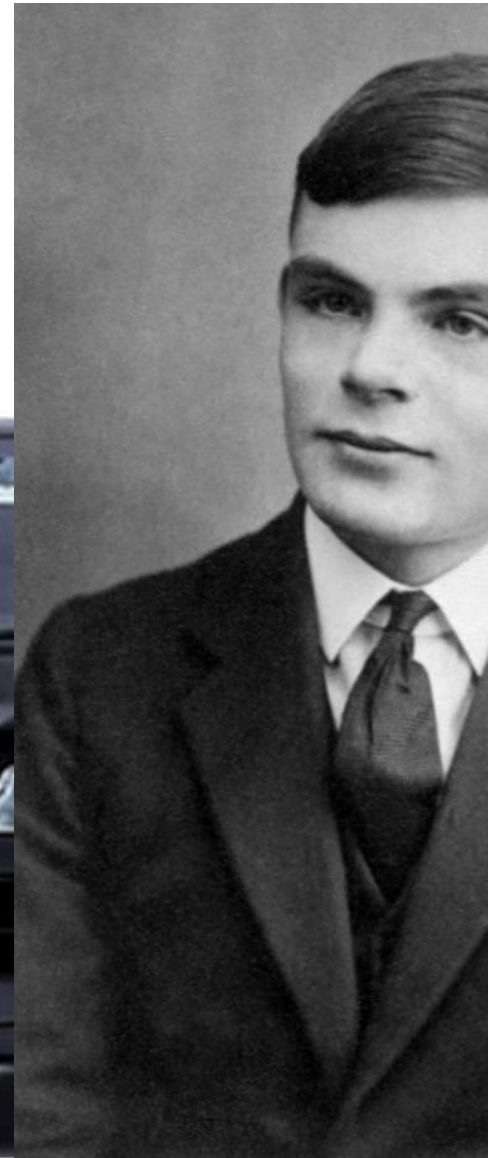


GPT-2
(1542M)

Source of image: <https://huaban.com/pins/1714071707/>

The models are become larger and larger ...

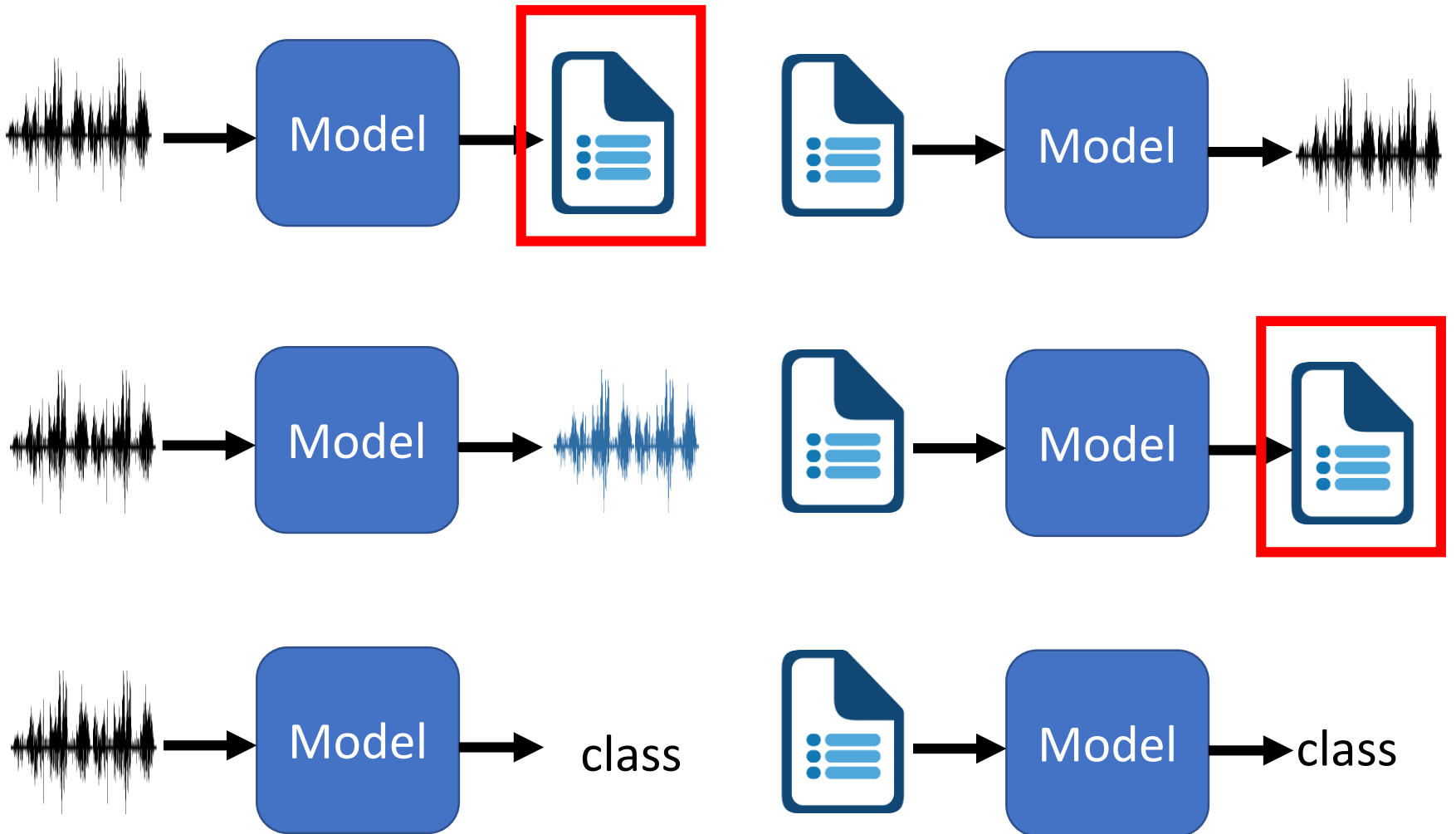
Turing NLG
(17G)



Megatron (8G)

T5 (11G)

One slide for this course




Text Generation

Autoregressive

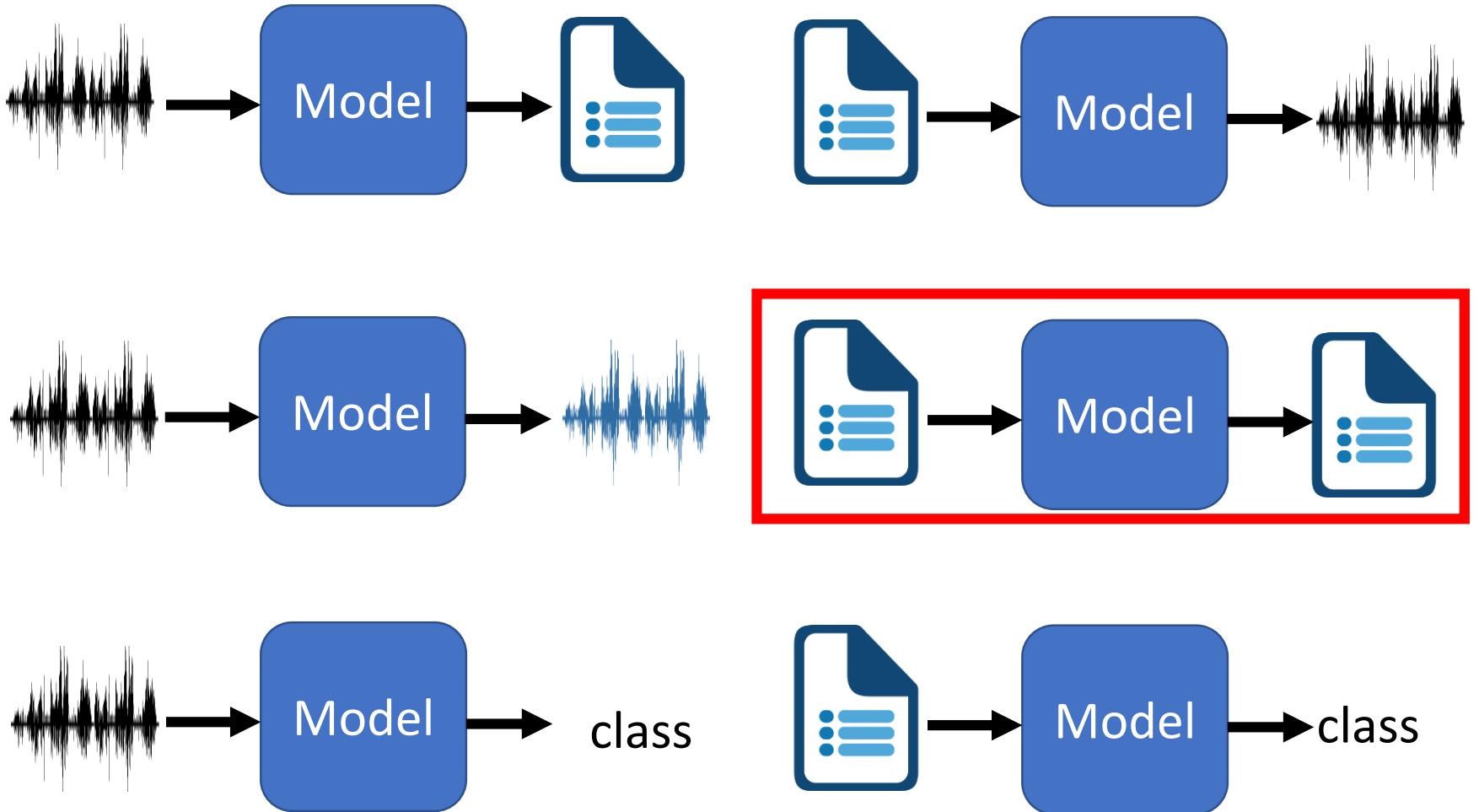
I → have → a → dream

Non-autoregressive

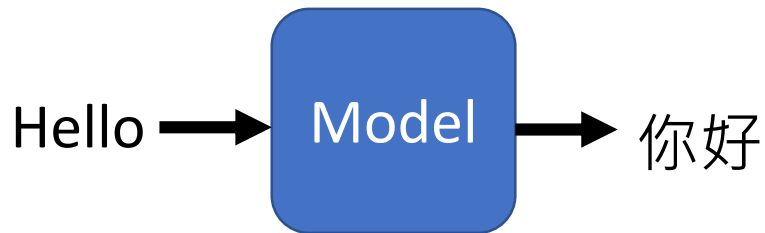
I → have → a dream

A curved arrow starts at the word 'dream' and points back to the word 'I', indicating a dependency between the end and the beginning of the sentence.

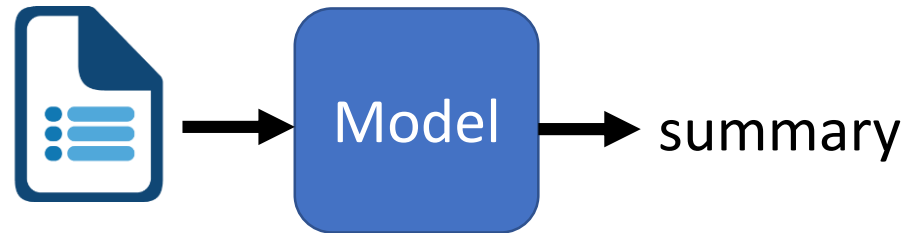
One slide for this course



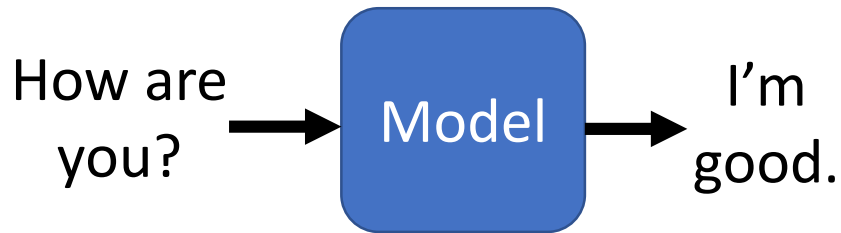
So many applications ...



Translation



Summarization



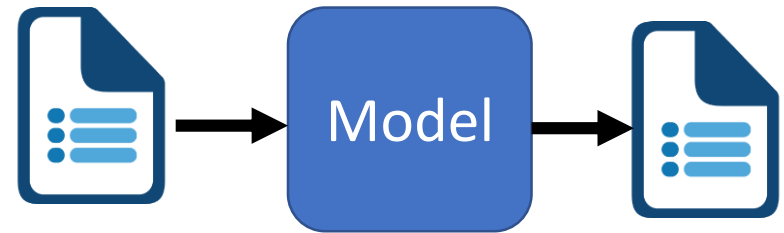
Chat-bot



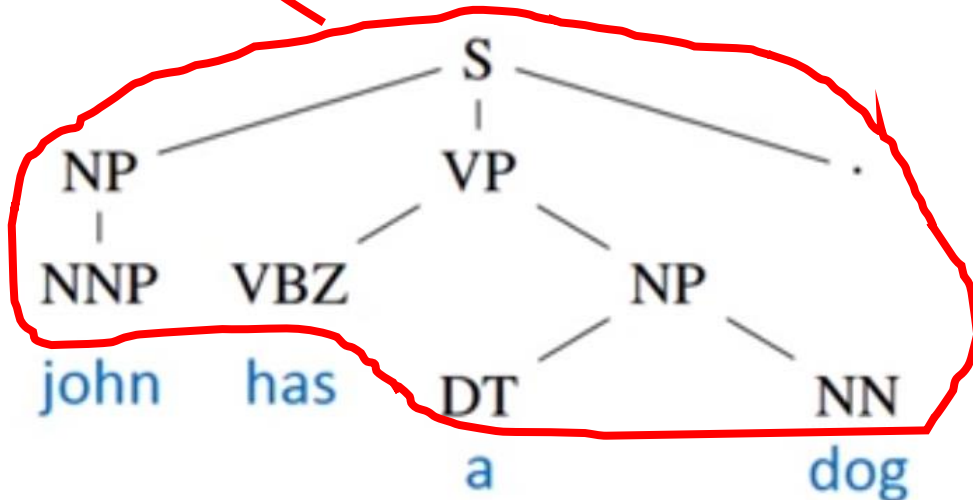
Question Answering

So many applications ...

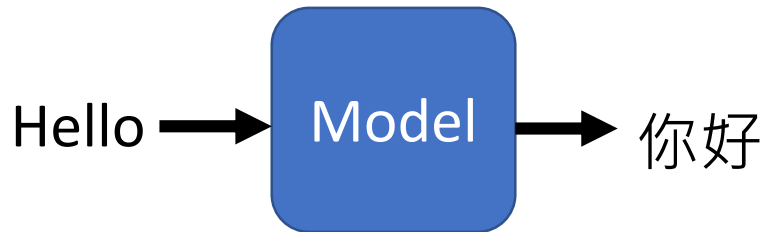
- Even syntactic parsing ...



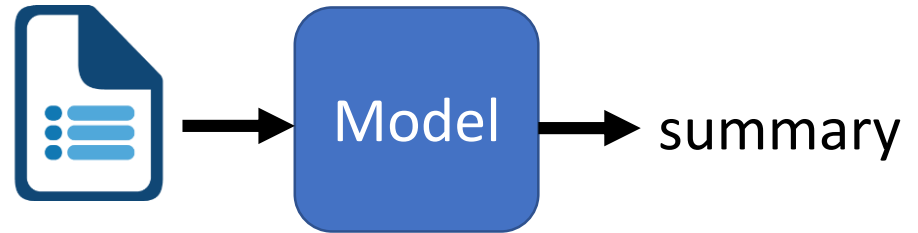
(S (NP NNP)_{NP} (VP VBZ (NP DT NN)_{NP})_{VP} .)_S



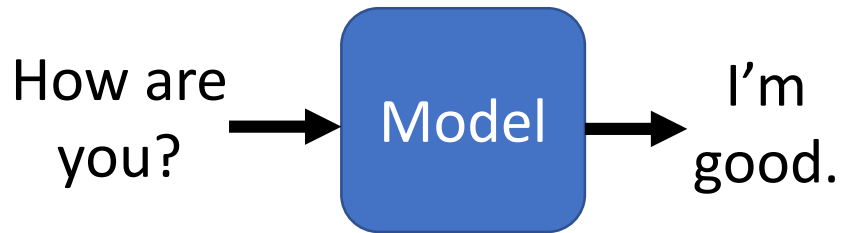
So many applications ...



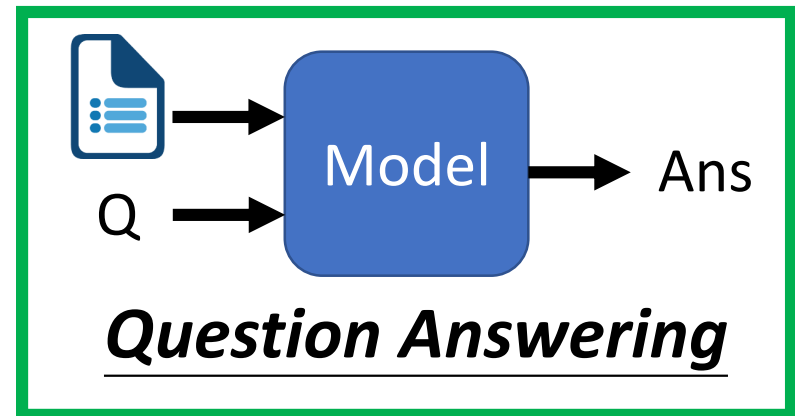
Translation



Summarization



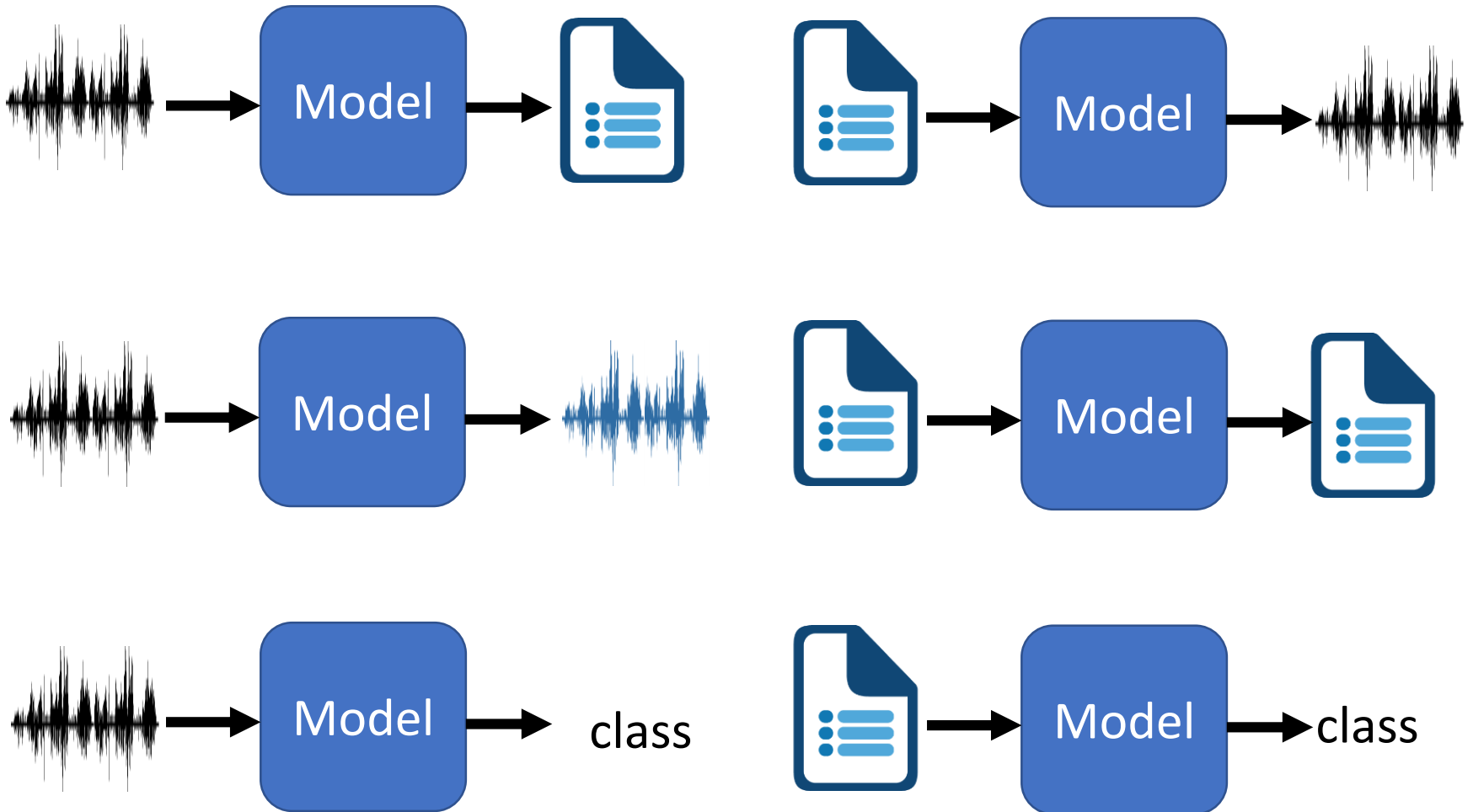
Chat-bot



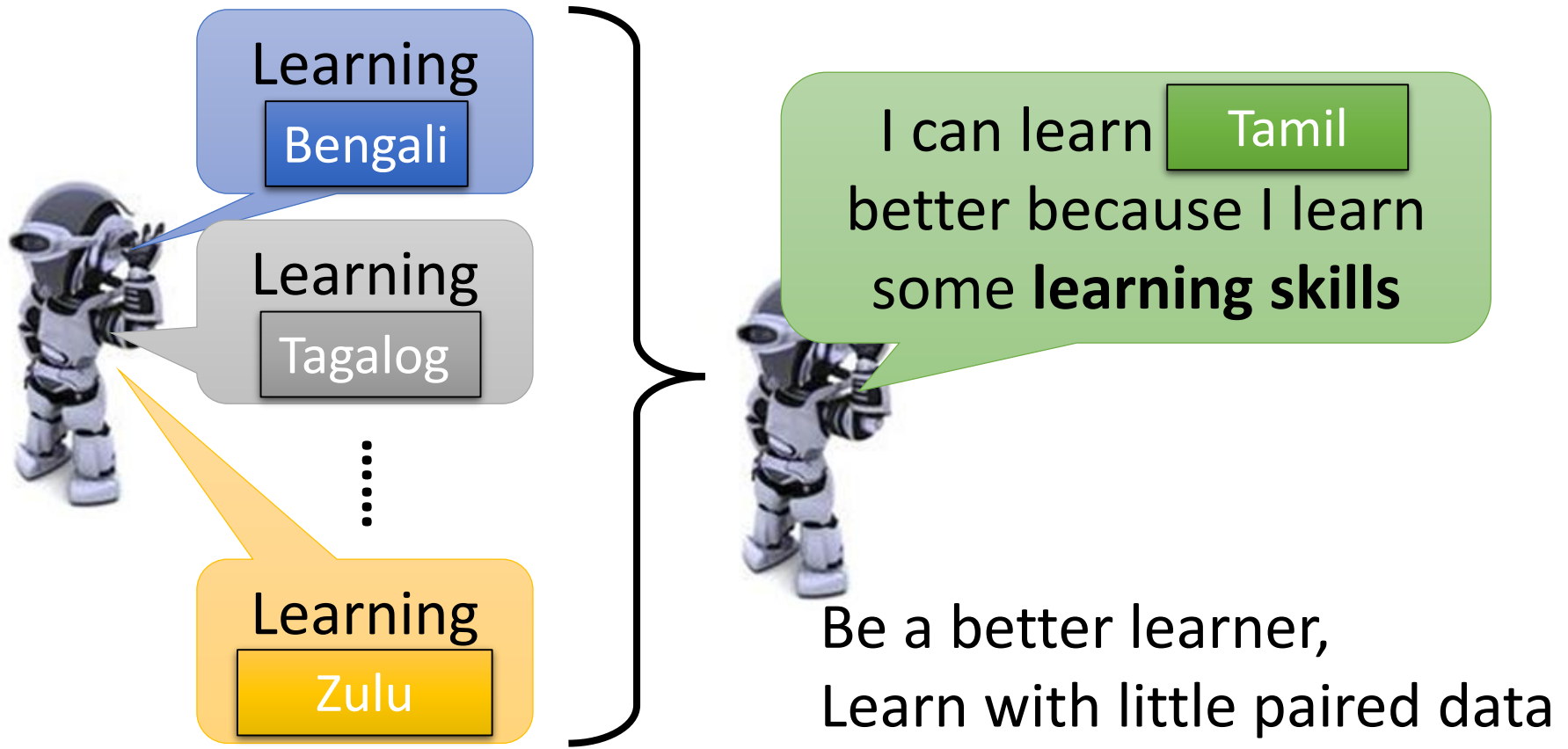
Question Answering

I will not go through all the applications because you will feel bored.

There is more



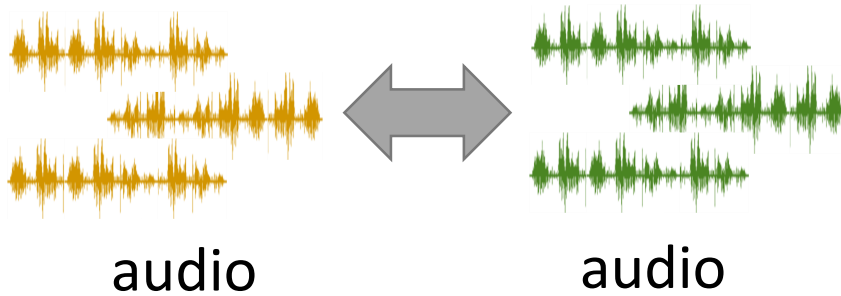
Meta learning = Learn to learn



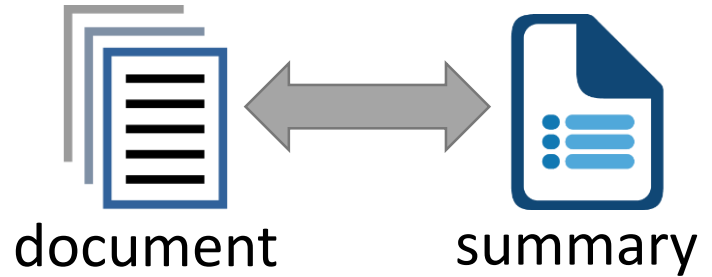
Learning from Unpaired Data



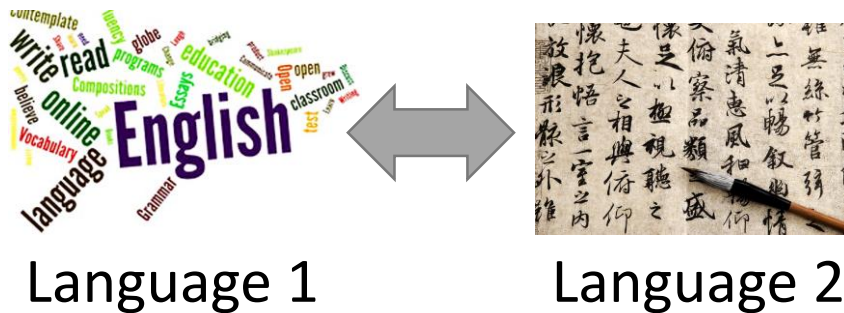
Image Style Transfer



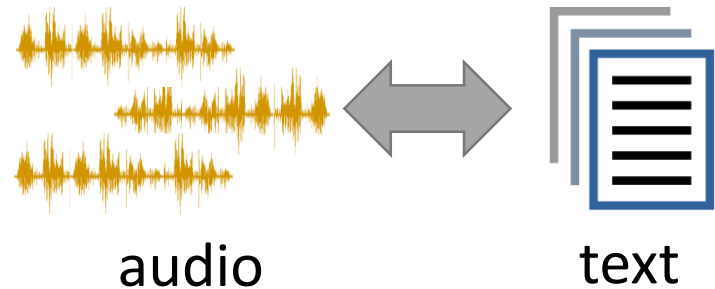
Voice Conversion



Summarization

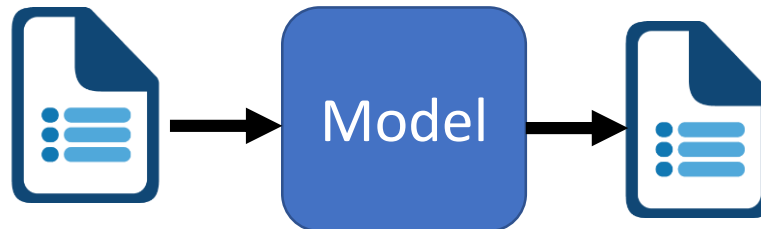
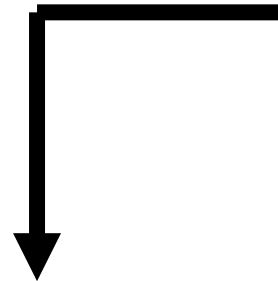
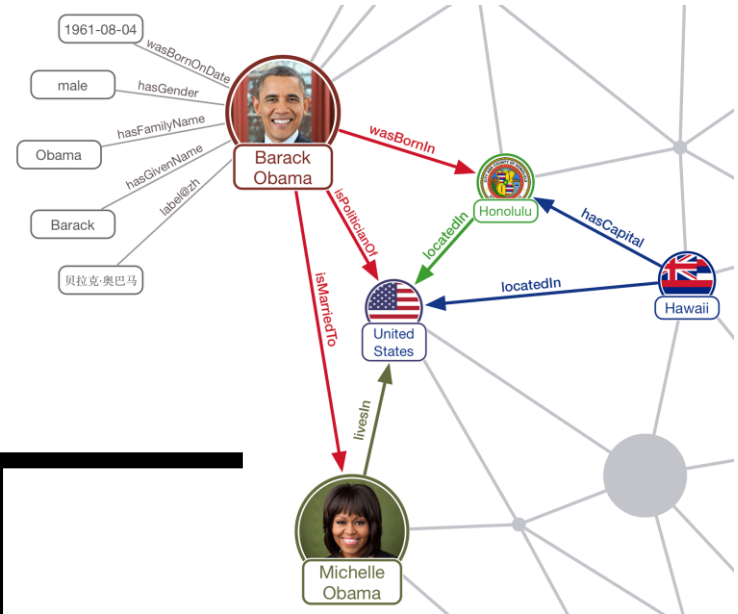


Unsupervised Translation



Speech Recognition

Knowledge Graph



Adversarial Attack

- Speech

- Anti-spoofing system (detecting synthesized speech) is easy to fool. [Liu, et al., ASRU, 2019]
- Speech recognition is easy to fool. [Lea Schonherr, et al., NDSS, 2019]

- NLP

Question: Why did he walk?

For exercise, Tesla walked between 8 to 10 miles per day. He squished his toes one hundred times for each foot every night, saying that it stimulated his brain cells.

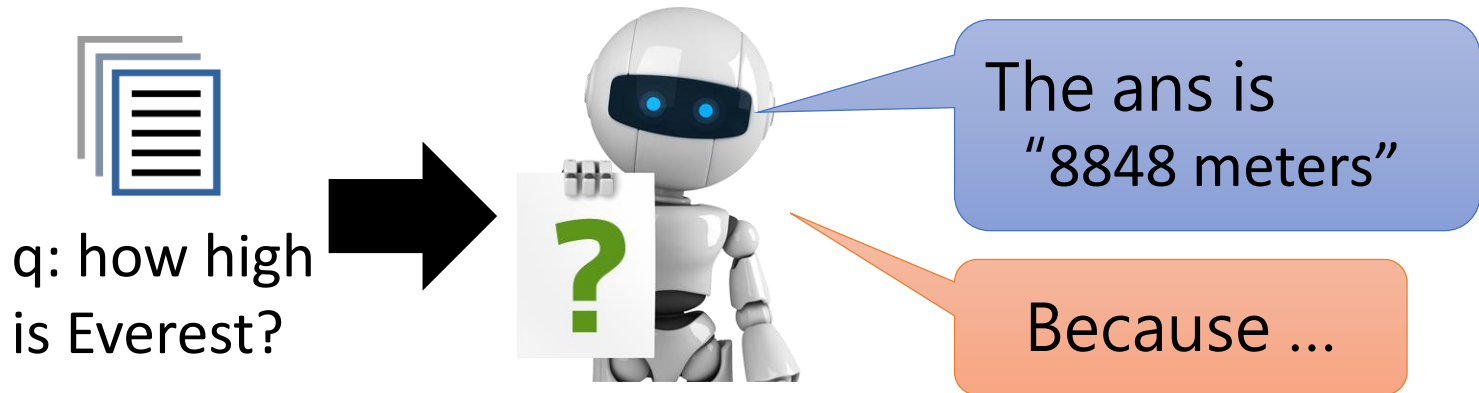
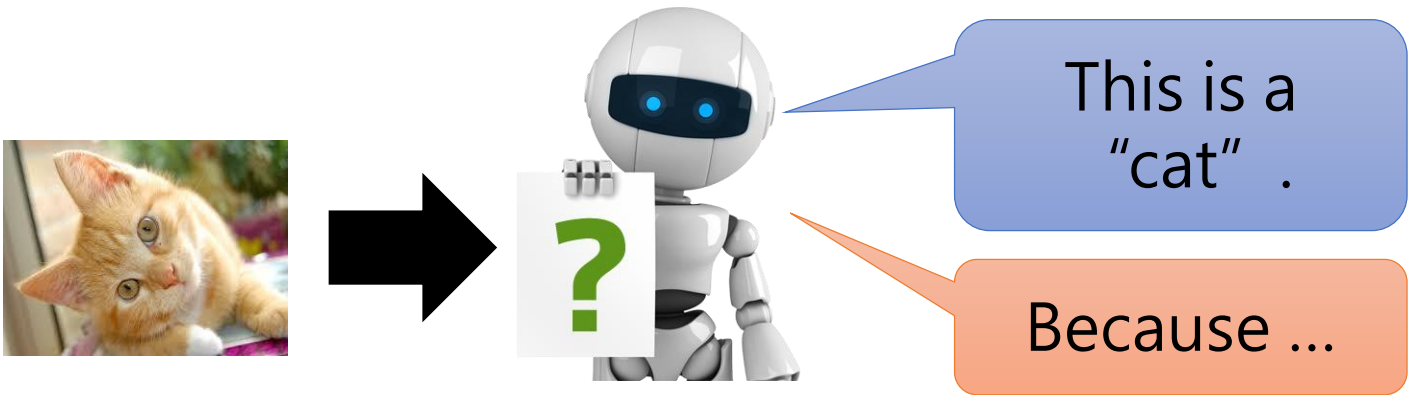
exercise

Question: Why did the university see a drop in applicants?

In the early 1950s, student applications declined as a result of increasing crime and poverty in the Hyde Park neighborhood. In response, the university became a

crime and poverty

Explainable AI



That's all for this course

