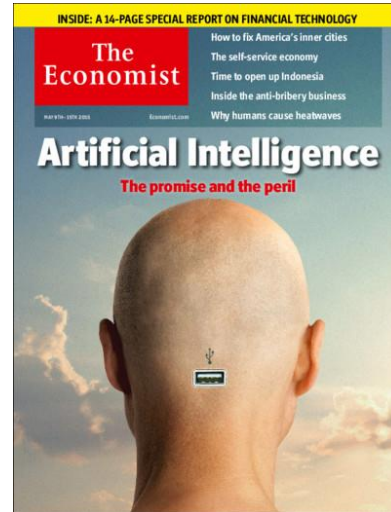# A Brief Intro to Machine Learning

Andres Masegosa
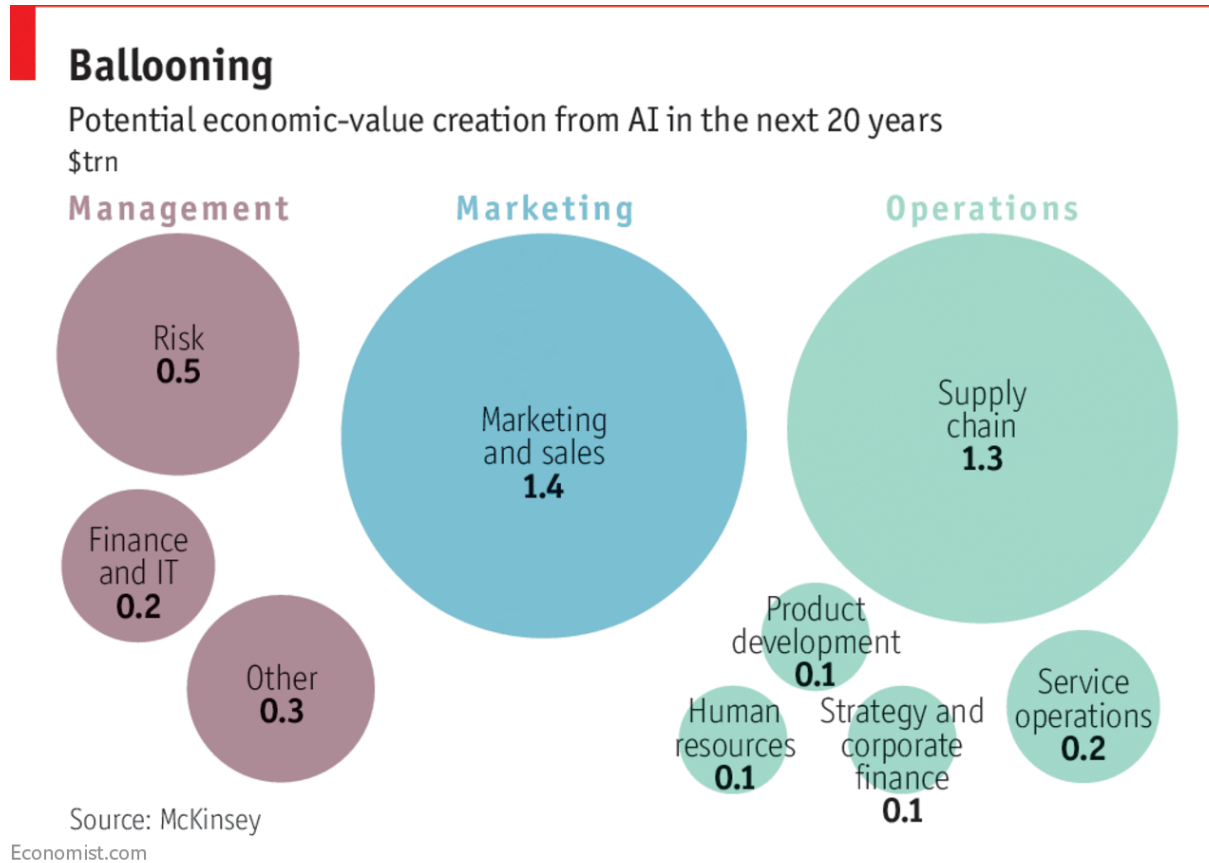
AALBORG UNIVERSITY

Machine Learning Impact/Hype
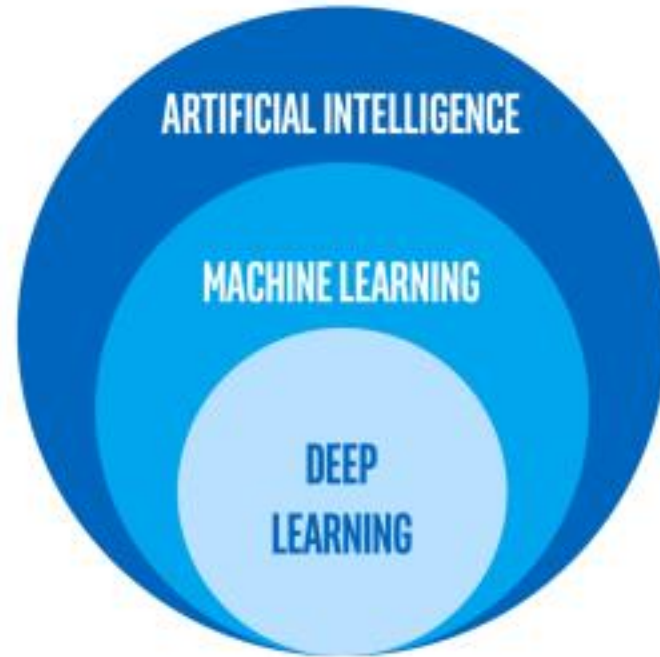
Machine Learning Impact/ Hype

Ballooning

Potential economic-value creation from AI in the next 20 years
$trn

Management — Risk 0.5; Finance and IT 0.2; Other 0.3

Marketing — Marketing and sales 1.4

Operations — Supply chain 1.3; Product development 0.1; Human resources 0.1; Strategy and corporate finance 0.1; Service operations 0.2

Source: McKinsey

Economist.com

Machine Learning Impact/ Hype

High Level Overview

Machine Learning Applications
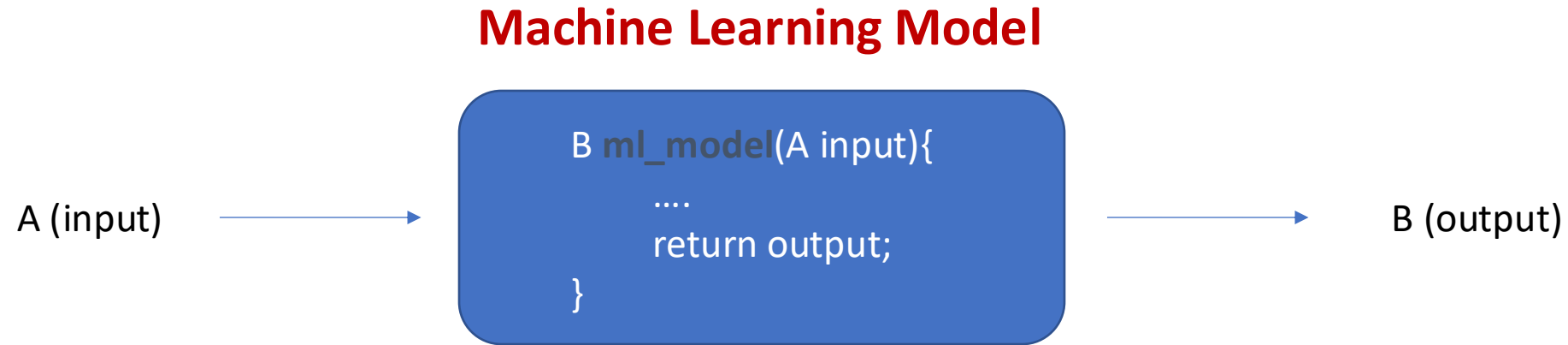
Machine Learning Model

A (input) →

B **ml_model**(A input){

    ….
    return output;

}

→ B (output)

Machine Learning Applications

Machine Learning Model

A (input) →

→ B (output)

Machine Learning Application

Spam Filter

Email (String) →

```
B ml_model(A input){
    ….
    return output;
}
```

→ Spam (Boolean)
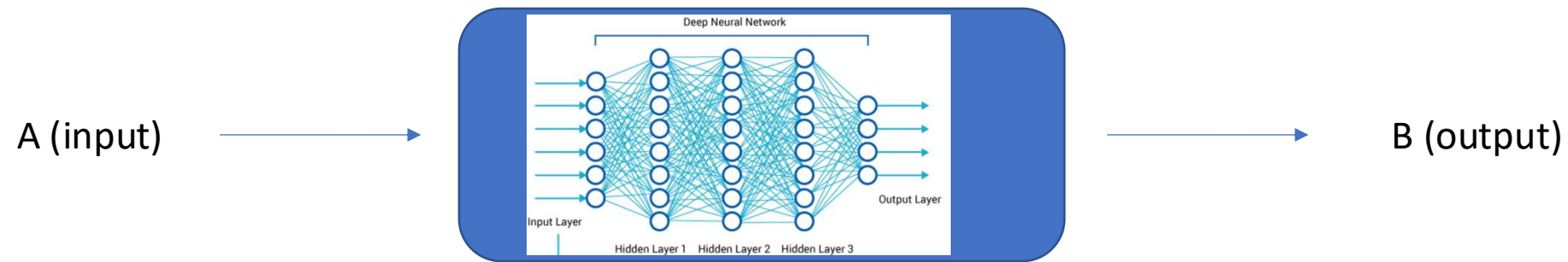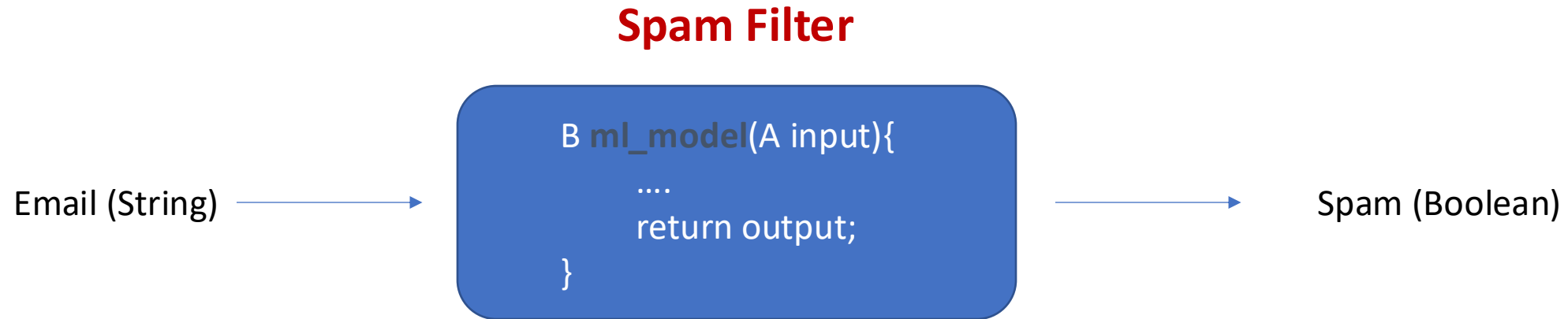
Machine Learning Application
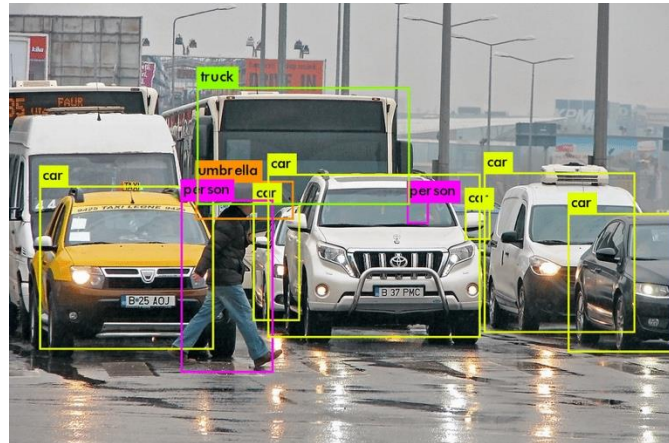
Object Dection

Image (int[][][]) →

```
B ml_model(A input){
    ....
    return output;
}
```

→ Objects (List<String>)

Machine Learning Application

Speech Recognition

Audio (doube[]) →

```
B ml_model(A input){
    ….
        return output;
}
```

→ Transcription (String)

Machine Learning Application

Automatic Translation

Chinese (String) →

```
B ml_model(A input){
    ….
    return output;
}
```

→ English (String)
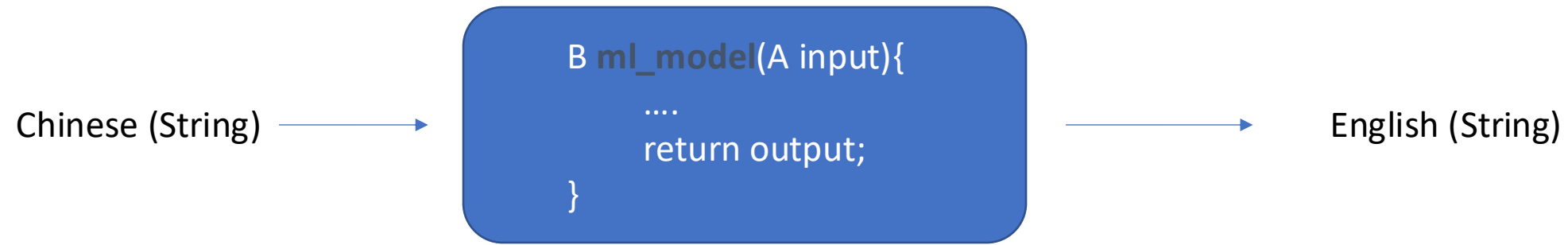
Machine Learning Application

Web Search

Query (String) →

```
B ml_model(A input){
    ....
    return output;
}
```
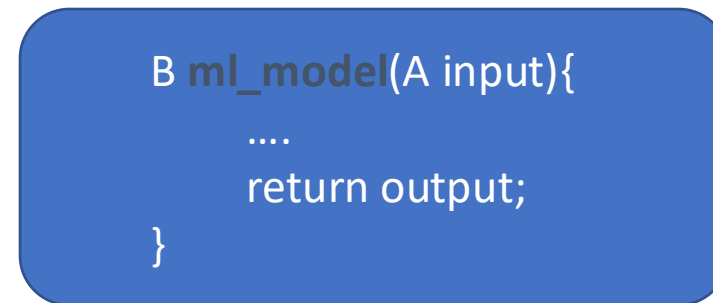
→ Web Links (List<String>)
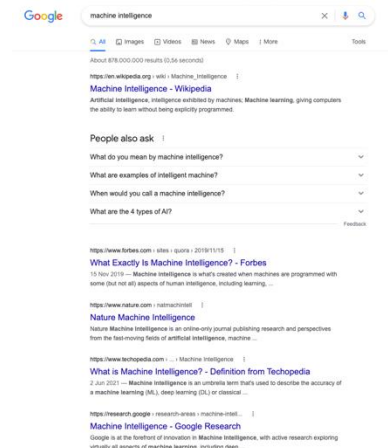
Machine Learning Application

Ad Placement

(Ad, Query, User) →

```
B ml_model(A input){
    ....
        return output;
}
```

→ Click? (boolean)
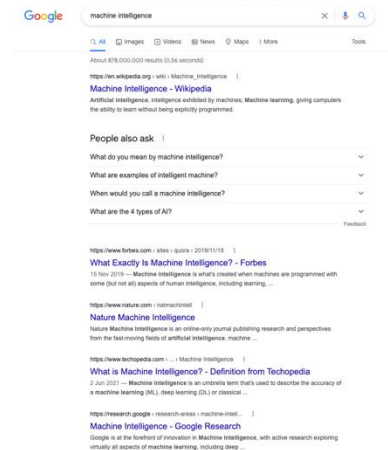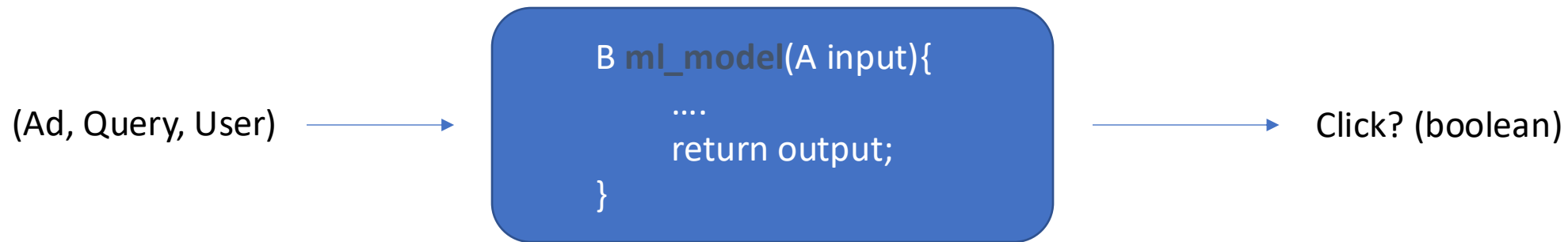
Machine Learning Application

Self-Driving Cars

(Cameras, LiDar, Sensors, Speed, Brakes, Wheel, etc)

```
B ml_model(A input){
    ....
    return output;
}
```

Car Action (Wheel + Pedals)

Machine Learning Application

Image Generation

Prompt (Text) →

```
B ml_model(A input){
    ….
    return output;
}
```

→ Image (int[][][])

https://imagen.research.google/

https://parti.research.google/

https://lexica.art/

Machine Learning Application

Protein Structure

Gene Sequence (int[])

Input sequence

```
B ml_model(A input){
    ....
        return output;
}
```

Protein Structure
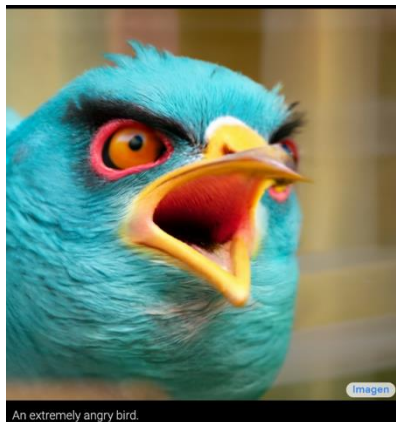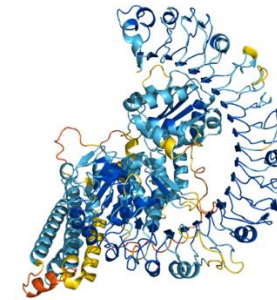
Article | Published: 15 January 2020

Improved protein structure prediction using potentials from deep learning

Andrew W. Senior ✉, Richard Evans, John Jumper, James Kirkpatrick, Laurent Sifre, Tim Green, Chongli Qin, Augustin Žídek, Alexander W. R. Nelson, Alex Bridgland, Hugo Penedones, Stig Petersen, Karen Simonyan, Steve Crossan, Pushmeet Kohli, David T. Jones, David Silver, Koray Kavukcuoglu & Demis Hassabis

Nature **577**, 706–710 (2020) | Cite this article

139k Accesses | **1120** Citations | **659** Altmetric | Metrics

Machine Learning Application

LLMs

Word Sequence (String[])

"There was once upon a"

```
B ml_model(A input){
    ….
    return output;
}
```

Next Word (String)

"time"
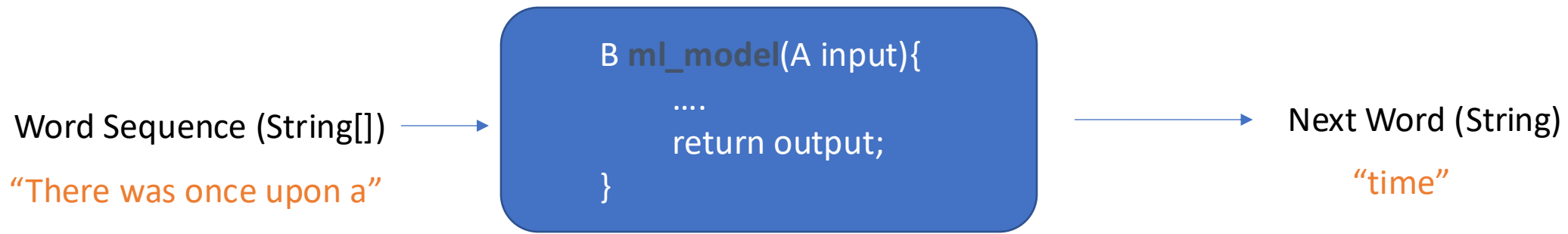
Exercise (5 min)

ML Application???

B **ml_model**(A input){

    ....

    return output;

}

Input?? →

→ Output??
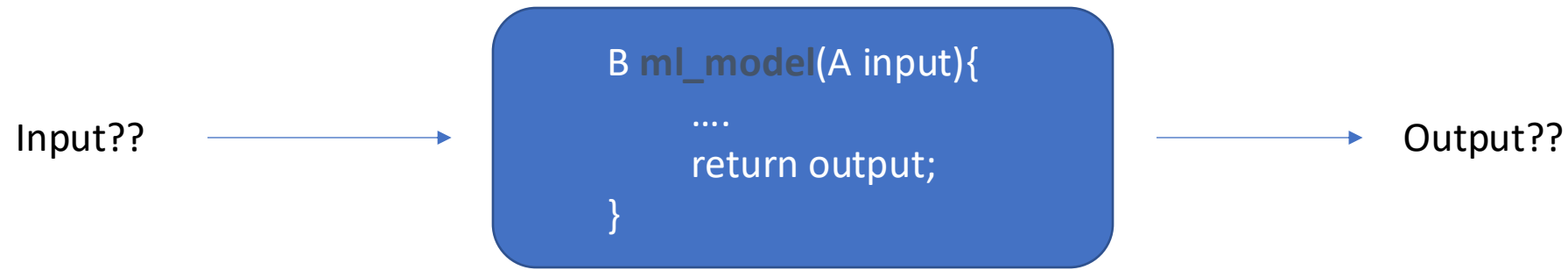
Think in applications that can be framed in this way

BREAK

Machine Learning Algorithms
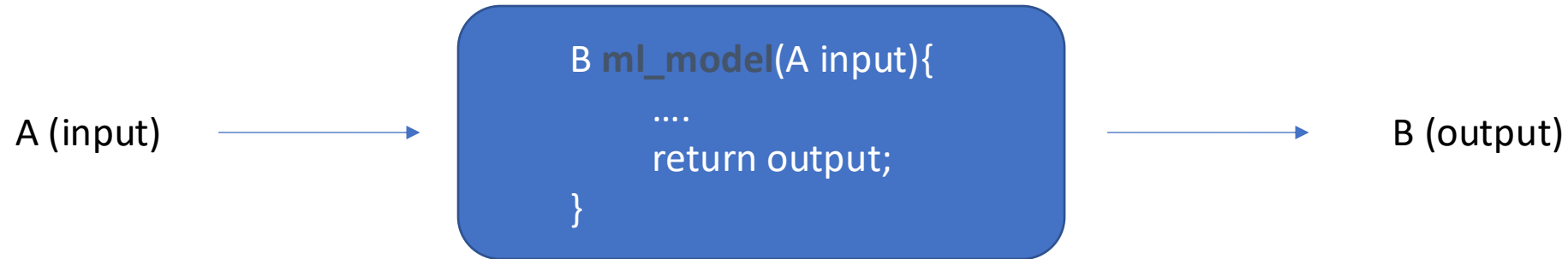
Machine Learning Model

A (input) →

```
B ml_model(A input){
    ....
        return output;
}
```

→ B (output)

Machine Learning Algorithms

Machine Learning Model

A (input) →

```
B ml_model(A input){
    ….
        return output;
}
```

→ B (output)

- -
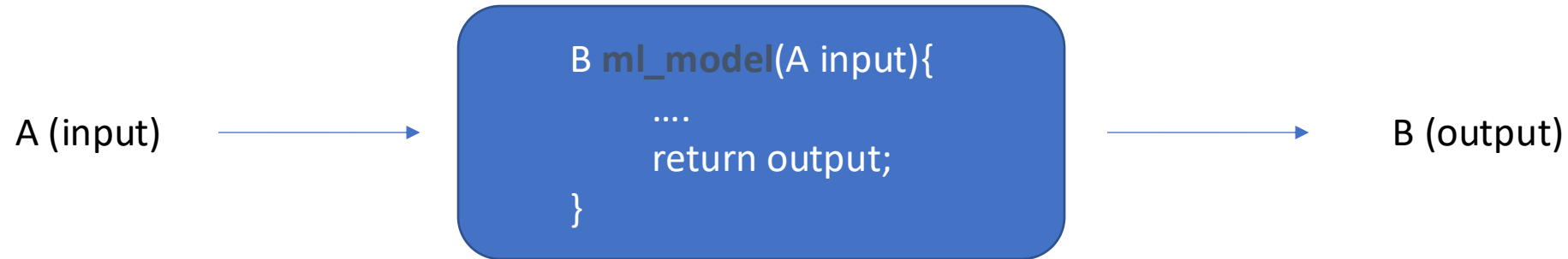
Data Set

List<(A,B)> →

ML Algorithm

→ **Machine Learning Model**

```
B ml_model(A input){
    ….
        return output;
}
```

Data is the Fuel of Machine Learning

Data Set

List<(A,B)>

ML Algorithm

Machine Learning Model

```
B ml_model(A input){
    ….
        return output;
}
```
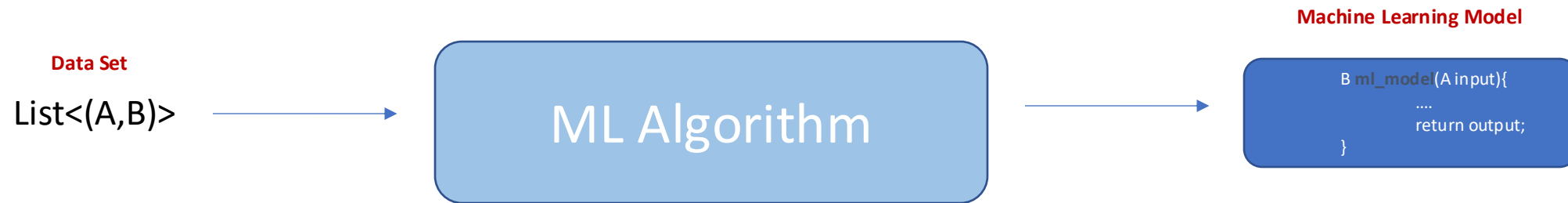
Definition by Tom Mitchell (1998):
Machine Learning is the study of algorithms that
- improve their performance P
- at some task T
- with experience E.
A well-defined learning task is given by $<P, T, E>$.

Data is the Fuel of Machine Learning

Andrew Ng: Artificial Intelligence is the New Electricity.
https://www.youtube.com/watch?v=21EiKfQYZXc&t=1206s

Data is the Fuel of Machine Learning

Automatic Translation

Machine Learning Model

Data Set

List<(A,B)>

ML Algorithm

```
B ml_model(A input){
    ....
    return output;
}
```

Danish Scandinavian	English
Hus	House
Barn	Child/Baby
Sæbe	Soap
Dag	Day
Gård(by)	Yard(City)
Vi	We
Sin	His/Her
Land(Jord)	Land(Earth)
Bund	Bottom
Modig/ Konge?	Brave/ King?
Knægt/Søn	Boy/ Knight?/Son
Bære	Bear
Bære(Køre)	Carry
Sjæl	Soul
Stellar	Stellar

Data is the Fuel of Machine Learning

Web Search

Data Set

List<(A,B)>

ML Algorithm

Machine Learning Model

```
B ml_model(A input){
       ....
       return output;
}
```

Query	Clicked Web Link
World Cup	https://www.fifa.com/fifaplus/en/tournaments/mens/worldcup/qatar2022
aau software bachelor copenhagen	https://www.aau.dk/uddannelser/bachelor/software
python matrix multiplication	https://numpy.org/doc/stable/reference/generated/numpy.matmul.html
....	....
....	....
....	....
....	....

Data is the Fuel of Machine Learning

Web Search

Machine Learning Model

Data Set

List<(A,B)>

ML Algorithm

```
B ml_model(A input){
        ....
        return output;
}
```

Query	Clicked Web Link
World Cup	https://www.fifa.com/fifaplus/en/tournaments/mens/worldcup/qatar2022
aau software bachelor copenhagen	https://www.aau.dk/uddannelser/bachelor/software
python matrix multiplication	https://numpy.org/doc/stable/reference/generated/numpy.matmul.html
....	....
....	....
....	....
....	....

How is Google collecting this data set?

Data is the Fuel of Machine Learning

Self-Driving Car

Machine Learning Model

Data Set

List<(A,B)>

ML Algorithm

```
B ml_model(A input){
    ....
    return output;
}
```

A = (Cameras, LiDar, Sensors,Speed, Brakes, Wheel, etc)

B = Car Action (Wheel + Pedals)

Data is the Fuel of Machine Learning

Self-Driving Car

Machine Learning Model

Data Set

List<(A,B)>

A = (Cameras, LiDar, Sensors,Speed, Brakes, Wheel, etc)

B = Car Action (Wheel + Pedals)

ML Algorithm

```
B ml_model(A input){
    ....
    return output;
}
```

How is Tesla collecting this data set?

Data is the Fuel of Machine Learning

Self-Driving Car

Machine Learning Model

Data Set

List<(A,B)>

ML Algorithm

B **ml_model**(A input){
.....
return output;
}

A = (Cameras, LiDar, Sensors,Speed, Brakes, Wheel, etc)

B = Car Action (Wheel + Pedals)

Estimated Tesla Autopilot Miles @lexfridman

Projected Autopilot Miles:
5,134,585,459
Date: 2021-01-01

Projected Autopilot Miles on Hardware 1:
994,077,265
Projected Autopilot Miles on Hardware 2/3:
4,140,508,193

Current Estimated Autopilot Miles:
3,324,020,246
Date: 2020-04-22

Current Autopilot Miles on Hardware 1:
852,461,468
Current Autopilot Miles on Hardware 2/3:
2,471,558,778

These two figures explain a lot...

Andrew Ng: Artificial Intelligence is the New Electricity.
https://www.youtube.com/watch?v=21EiKfQYZXc&t=1206s
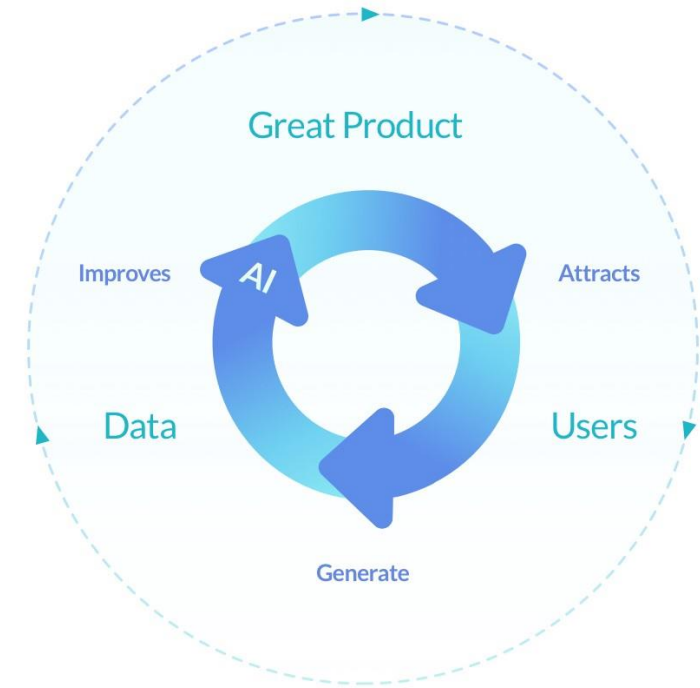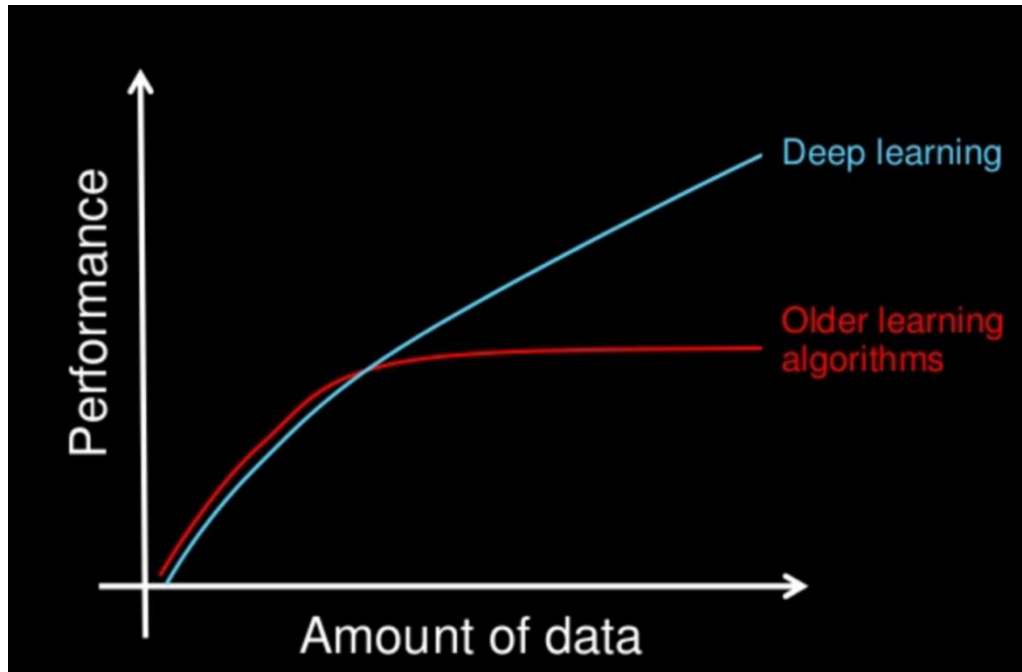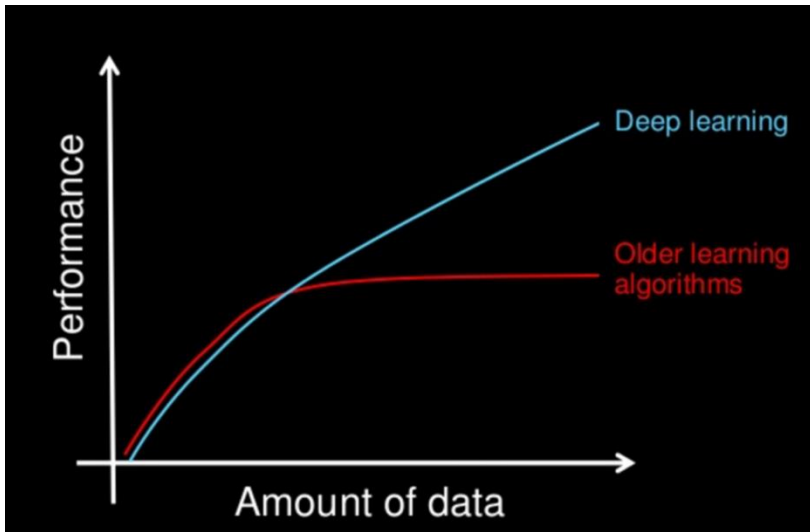
Is Google a Monopoly? Exercise (10 min)

- Dicuss in Groups wether Google is a Monopoly?

- Facts:
  - Google pays billions to be the default search engine, including $20bn to Apple in 2022.
  - 50% of US searches happen on default platforms where Google has contracts.

U.S. Said to Consider a Breakup of Google to Address Search Monopoly

The Justice Department and state attorneys general are discussing various scenarios to remedy Google's dominance in online search, including a breakup of the company.

Google's chief executive, Sundar Pichai, at a company event last year. The tech giant acted illegally to maintain a monopoly in online search, a federal judge ruled this month. Jason Henry for The New York Times

Great Product
Improves
AI
Attracts
Data
Users
Generate
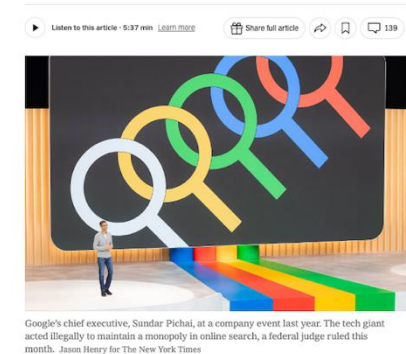
ML also needs a lot computer power....

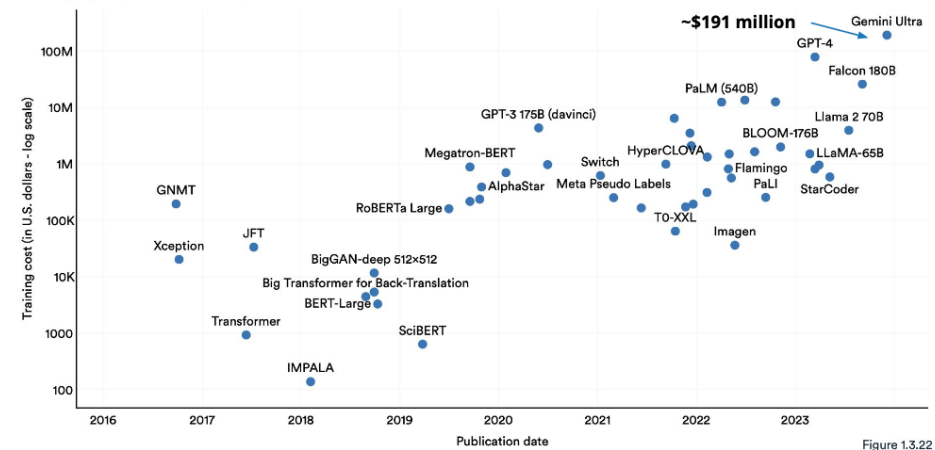Andrew Ng: Artificial Intelligence is the New Electricity.
https://www.youtube.com/watch?v=21EiKfQYZXc&t=1206s

HAI_2024_AI-Index-Report
https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_2024_AI-Index-Report.pdf

Exercise (5 min)

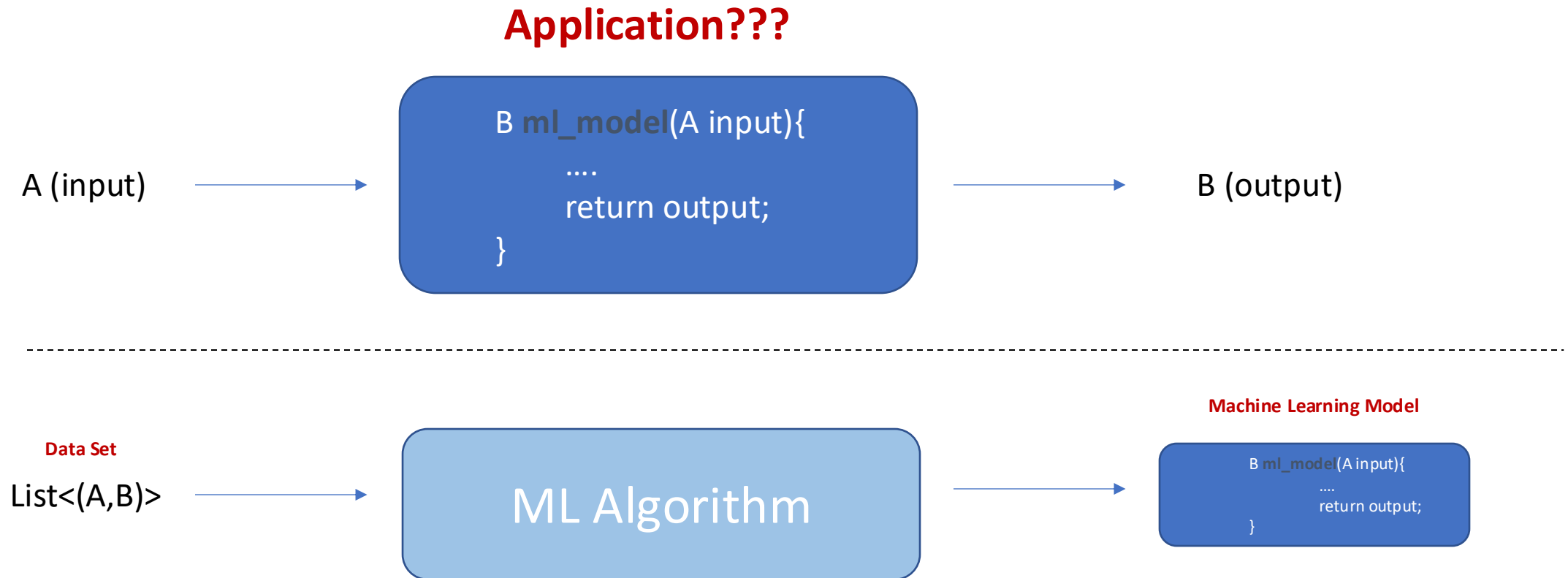**Application???**

A (input) →

B **ml_model**(A input){

    ….

       return output;

}

→ B (output)

Machine Learning Model

Data Set

List<(A,B)> → **ML Algorithm** →

B ml_mode (A input){

    ….

       return output;

}

How could you collect data for your application?