

# End-to-end Industrial Practice on Data Engineering and Machine Learning

葉信和 / Hsin-Ho Yeh  
Software Engineer / Funder / CEO @ 信誠金融科技  
[hsinho.yeh@footprint-ai.com](mailto:hsinho.yeh@footprint-ai.com)

# Download Slides

<https://reurl.cc/jRD3ED>



# About me

- 2020 - Present at 信誠金融科技
  - Shrimping: A data-sharing platform
    - <https://get-shrimping.footprint-ai.com>
  - Tintin: a machine learning platform for everyone
    - <https://get-tintin.footprint-ai.com>
  - KaFeiDo: machine learning platform for green economy
- 2016 - 2020 at Igloolnsure (16M+ in series A+ 2020)
  - Provide digital insurance for e-economic world
  - Funded in KUL, Headquartered in Singapore
  - First employee/ Engineering Lead / Regional Head/ Chief Engineer
- 2013 - 2016 at Studio Engineering @ hTC
  - Principal Engineer on Cloud Infrastructure Team
- 2009 - 2012 at IIS @ Academia Sinica
  - Computer vision, pattern recognition, and data mining
- CS@CCU, CS@NCKU alumni



# Agenda

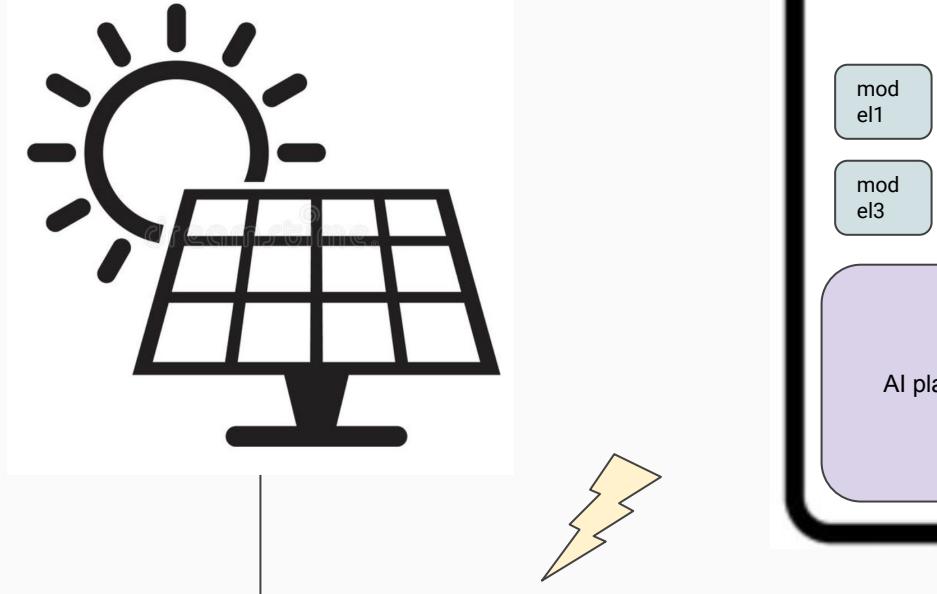
- Who we are and what we do.
- End-to-end data engineering and machine learning walkthrough.
- FAQ

- 無所不在的AI - 但需要把資料送回後端機房處理
  - Globally, data transmission networks consumed 260-340 TWh in 2020, or 1.1-1.4% of global electricity use. [1]
- MLDL模型效能特別好 - 但只適用於已知資料集
  - Is it reasonable to use copurs between 2010-2015 to predict what people is talking about in 2022?
  - Is it reasonable to train a car detector from 90s car dataset?
- 我們的電腦跑得很快 - 但需要機房的低溫設置避免熱當
  - In 2014, data centers in the U.S. consumed an estimated 70 billion kWh, representing about 1.8% of total U.S. electricity[2]

[1] <https://www.iea.org/reports/data-centres-and-data-transmission-networks>

[2] <https://www.techtarget.com/searchdatacenter/tip/How-much-energy-do-data-centers-consume>

# Ubiquitous AI platform (夢裡什麼都有?)

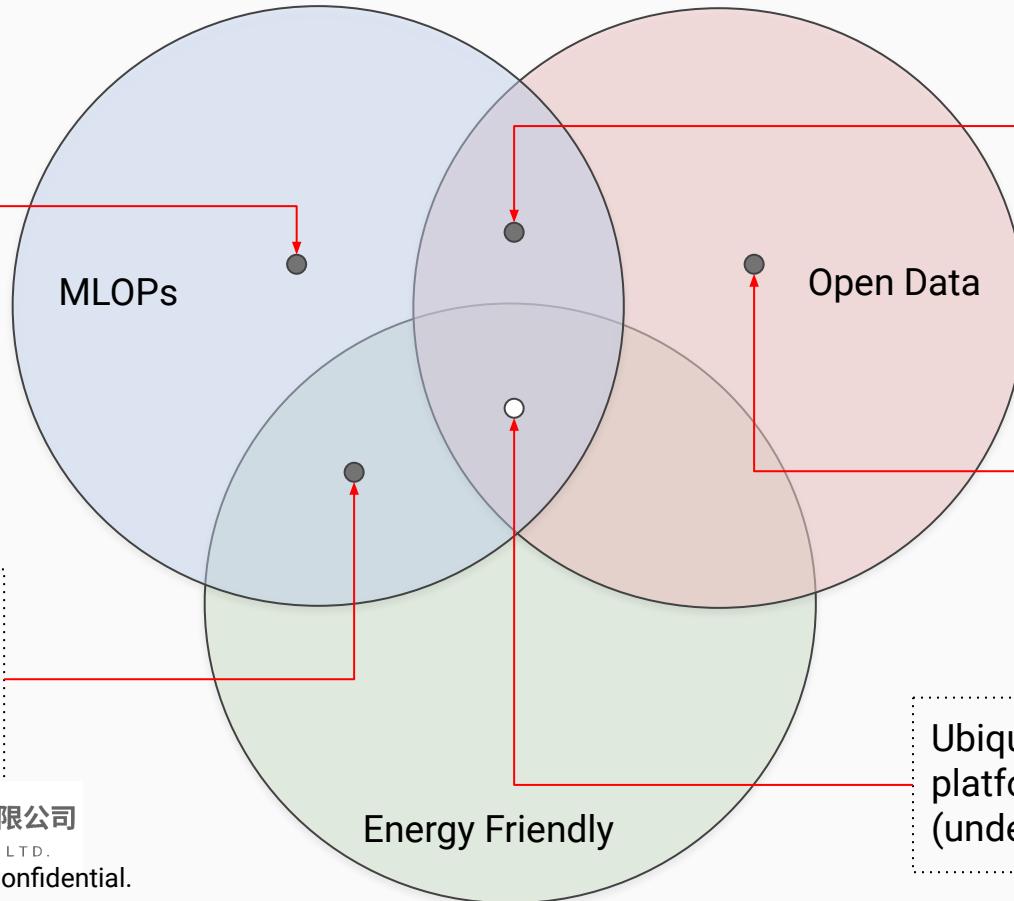


- Sustainability: Self-charging

- Energy-friendly: low power consumption
- Ubiquitous: device can be carried to anywhere, even network is not accessible.

# RoadMap

Tintin: Machine learning platform for everyone (2020-2021)



Grandturks: trustworthy AI platform for smart city (2022)

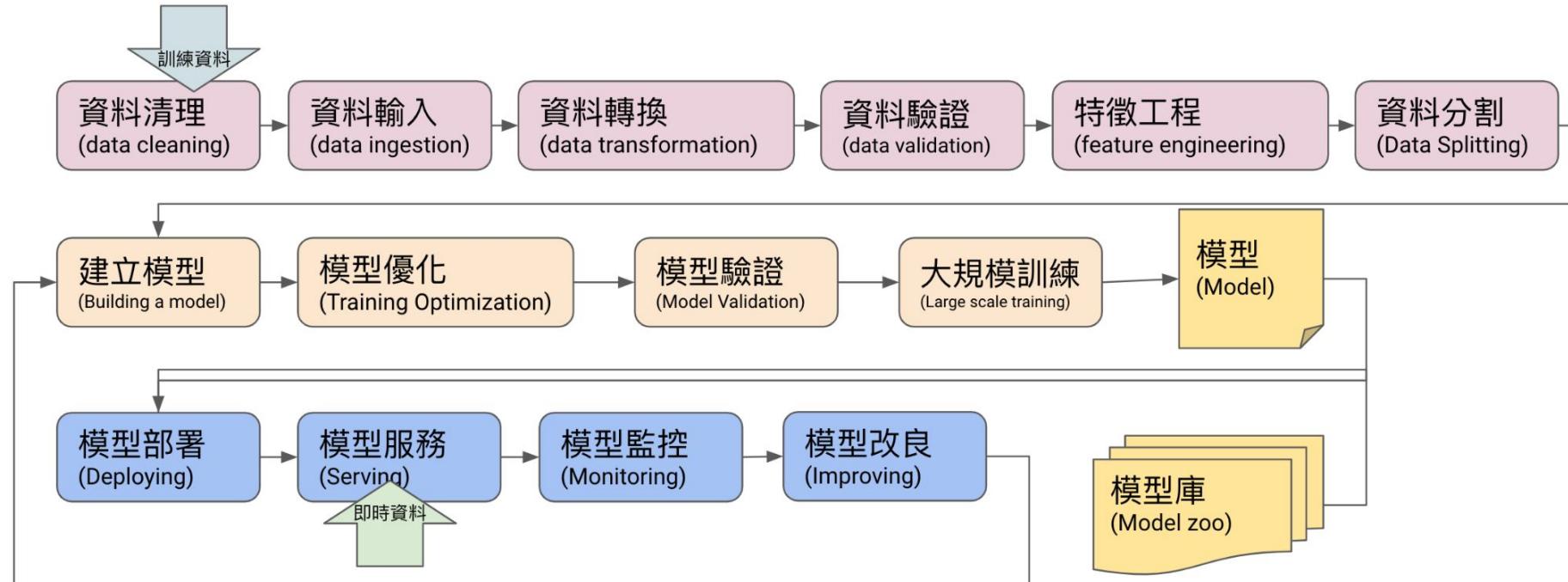
Shrimping: a data-sharing platform (2021)

KaFeDo: Machine learning platform for green economy (2022)

Ubiquitous AI platform (under developing)



# Real-world Machine Learning Application - End-to-End ML LifeCycle

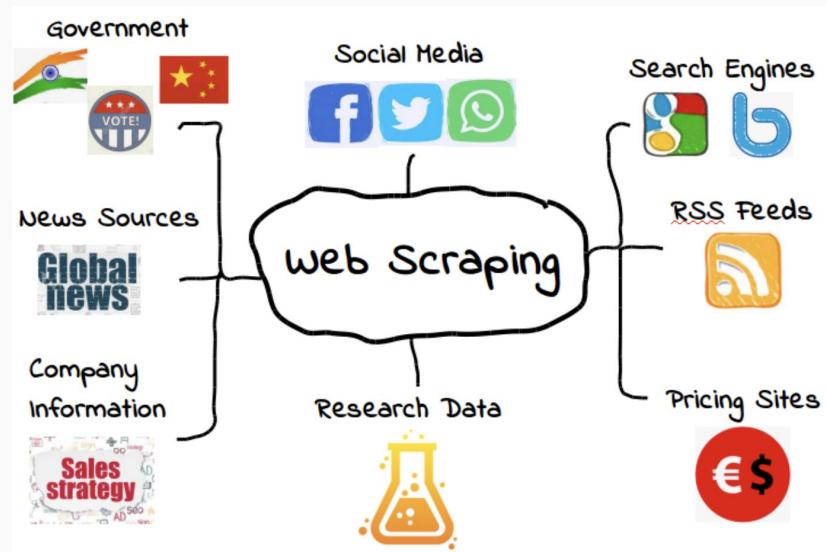


# Gathering Data Scraping With Selenium

Scraping is gathering information from public into your system for further analysis.

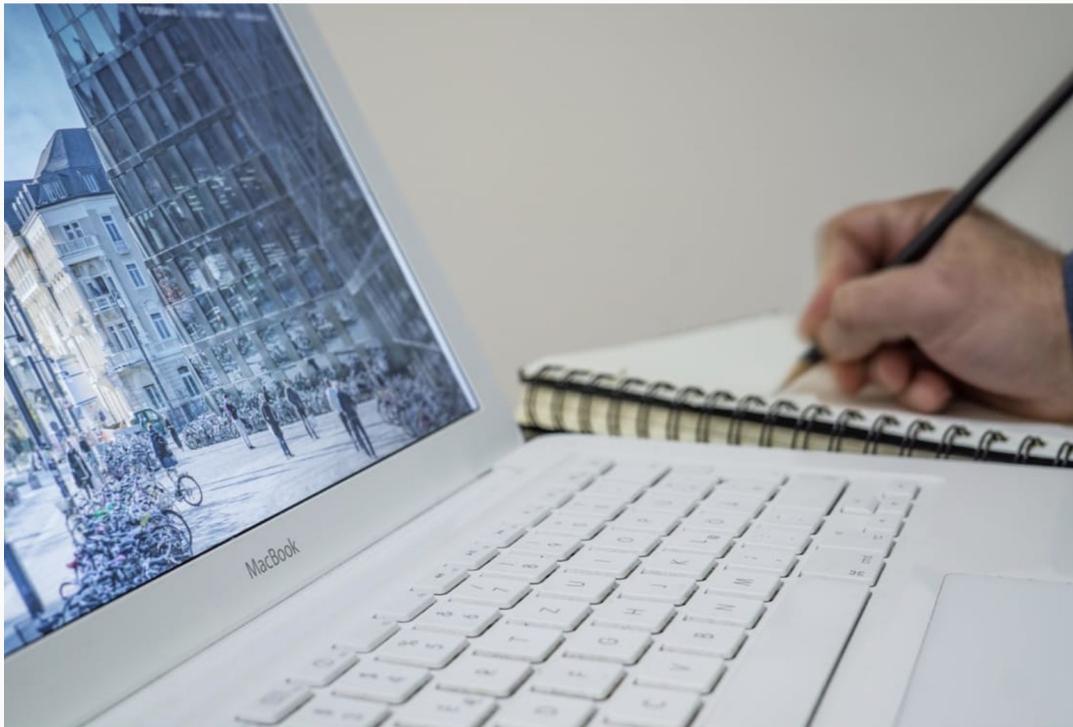
# What is Web Scraping?

- Web scraping is an **automation** process of collecting **structure data** from public websites.
- Common Use Cases including
  - Competitor price monitoring on E-commerce
  - News monitoring from social network



## Different ways of collecting data from websites (1/3)

- Manually (slow & slow & slow)



## Different ways of collecting data from websites (2/3)

- API Requests with Python(Stable & Fast but not always works...Why?
  - Not human-like interaction - anti-bot policy
  - Dynamic content
  - Too Frequently access

```
>>> import requests  
  
>>> r = requests.get('https://www.google.com.tw/')  
  
>>> print(r.status_code)  
200  
  
>>>print(r.text)  
<!doctype html><html itemscope="" itemtype="http://schema.org/WebPage"  
lang="zh-TW"><head><meta content="text/html; charset=UTF-8" http-equiv="Content-Type"><meta  
content="/images/branding/googleg/1x/googleg_standa....
```

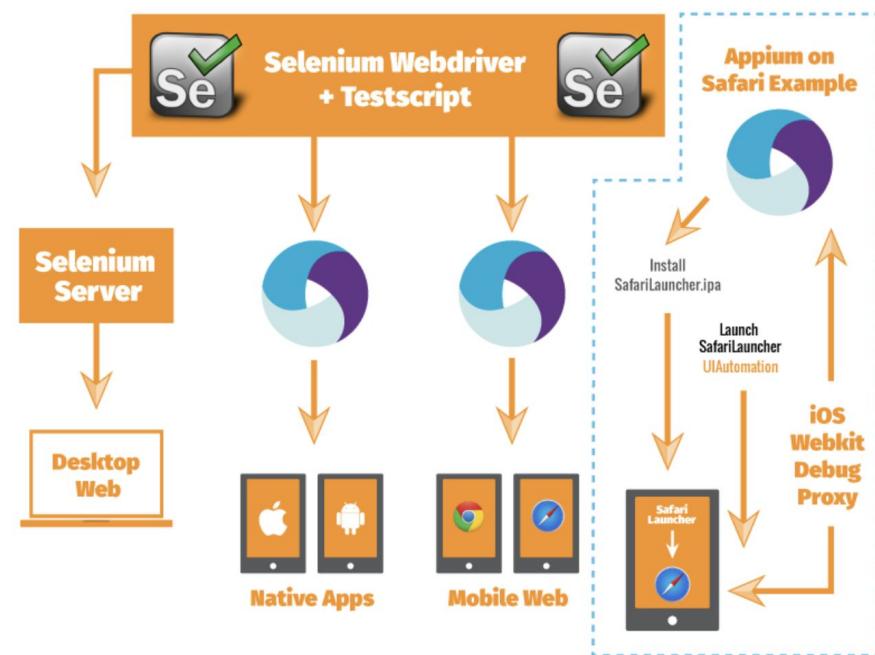
# Key fields sent by Browser

- Without this information (including cookie, referrer, user-agent fields), your program is easy to detect as a bot and get banned when you are accessing restricted resources.

```
x Headers Payload Preview Response Initiator Timing Cookies
▼ request headers
:authority: www.instagram.com
:method: POST
:path: /api/v1/feed/timeline/
:scheme: https
:accept: */*
:accept-encoding: gzip, deflate, br
:accept-language: en-US,en;q=0.9
:cache-control: no-cache
:content-length: 153
:content-type: application/x-www-form-urlencoded
cookie: mid=YpYwgAAEAAHaopTmodGr9VGA714p; ig_did=B8DC4143-652A-4F91-8CC7-7861EEE01533; ig_nrcb=1; ds_user_id=50697805654; csrf_token=sw4exlGoxUhWR2dK5yH; sessionid=50697805654%3ABXtwHsiI8CYm6Z%3A6%3AAYf-xPVGwmv536hXeB6WpCLJvBM5m65kDSHnTzHUqw
dnt: 1
origin: https://www.instagram.com
pragma: no-cache
referer: https://www.instagram.com/
sec-ch-prefers-color-scheme: light
sec-ch-ua: "Google Chrome";v="107", "Chromium";v="107", "Not=A?Brand";v="24"
sec-ch-ua-mobile: ?0
sec-ch-ua-platform: "macOS"
sec-fetch-dest: empty
sec-fetch-mode: cors
sec-fetch-site: same-origin
sec-gpc: 1
user-agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/107.0.0.0 Safari/537.36
viewport-width: 1680
```

## Different ways of collecting data from websites (3/3)

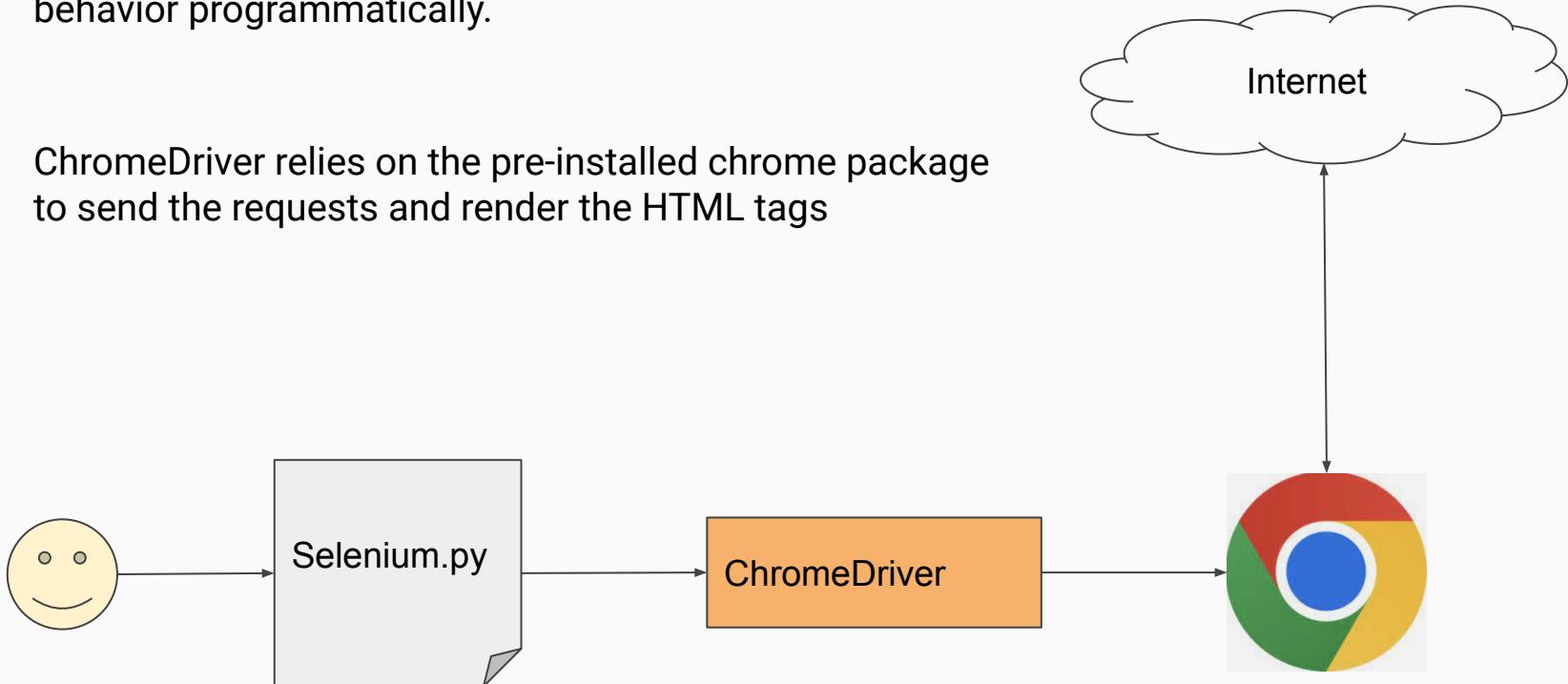
- UI Automation with Selenium
  - Selenium is an web automation tool which allows you to automate UI testing for different browsers via its webdriver.
  - Webdriver is used to control each browser's behavior making it more like real-human interaction.
- Because of its nature of automation and mimic human behavior, making it a good human-like web-scraping



Ref: <https://smartbear.com/blog/selenium-cross-browser-testing-on-mobile-devices/>

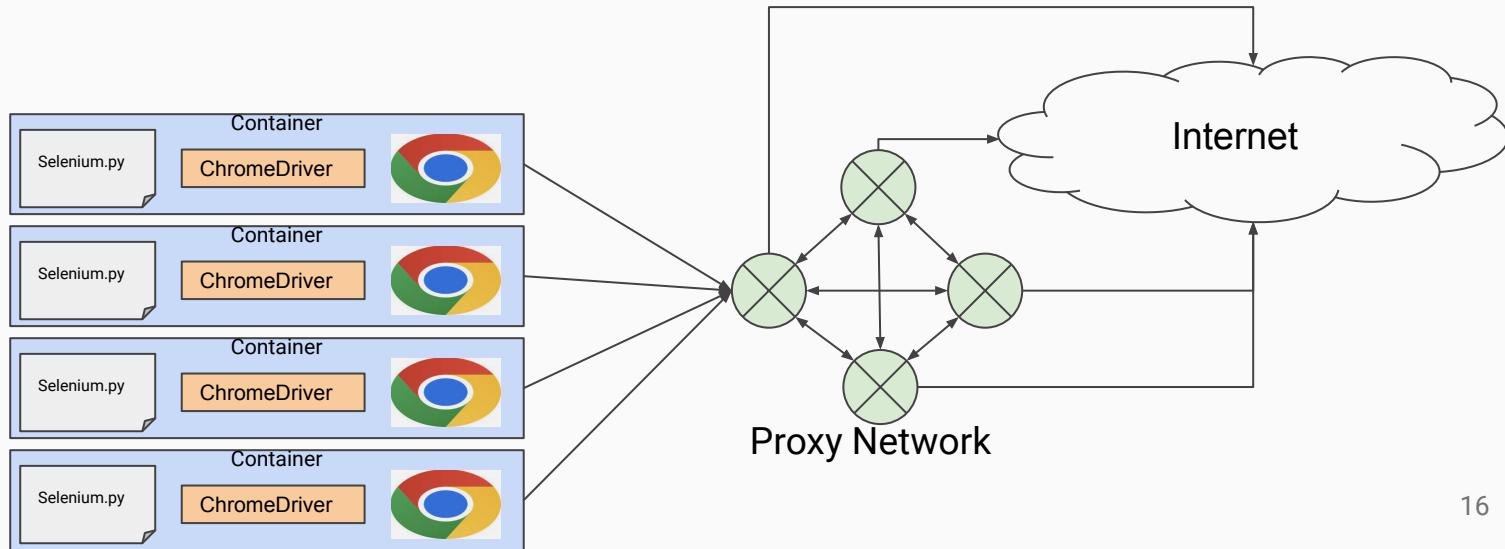
# How Selenium WebDriver Works?

- ChromeDriver provides an interface to communicate with the underlying chrome browser, allows you to control its behavior programmatically.
- ChromeDriver relies on the pre-installed chrome package to send the requests and render the HTML tags



# How Selenium WebDriver Works In Scale?

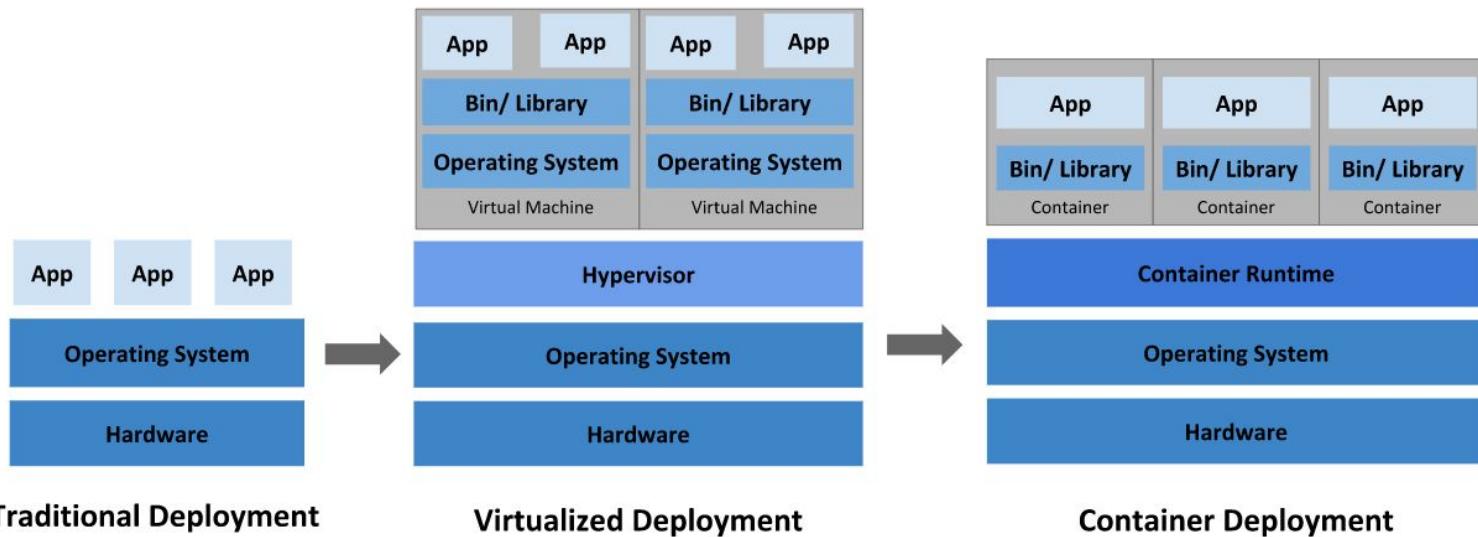
- Container provides a way to encapsulate your program into a single container image, allowing you to run it on any platform. (easy for scaling up/down)
- Proxy provides a good way to hide your source(IP / Identity) when you are accessing the public internet. The destination server won't know who is scraping their data.



# What is container?

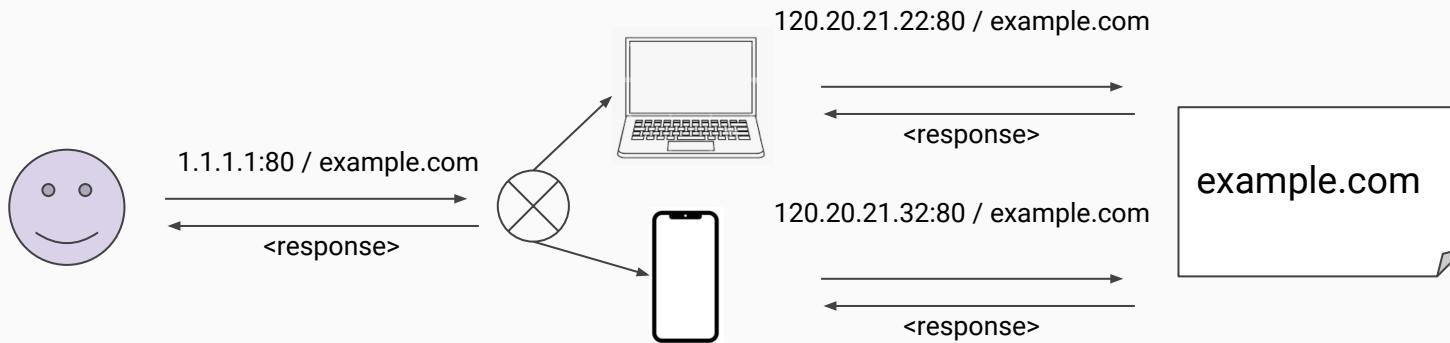
- Container

- Container Image = Application code + dependencies
- Runtime environment (cgroups, namespaces, env vars)



# How Proxy is working?

- Hide your tail (your actual ip address)
- Improve speed with content caching



# Selenium Examples

<https://github.com/FootprintAI/selenium-example>

The screenshot shows a GitHub repository page for 'FootprintAI / selenium-example'. The repository is public and has 1 branch and 0 tags. The main file list includes 'hsinhyeh remove', 'drivers', 'examples', 'Dockerfile', 'LICENSE', 'README.md', and 'chromedriver'. On the right side, there is a 'Clone' section with options for HTTPS, SSH, and GitHub CLI, along with a copy link. Below it are links for 'Open with GitHub Desktop' and 'Download ZIP'.

Code or jump to... Pull requests Issues Marketplace Explore

FootprintAI / **selenium-example** Public Edit Pins Unwa

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

main 1 branch 0 tags Go to file Add file Code

hsinhyeh remove

- drivers add headless as options
- examples remove
- Dockerfile add Dockerfile
- LICENSE Initial commit
- README.md Initial commit
- chromedriver add ig/goodinfo example

Clone

HTTPS SSH GitHub CLI

<https://github.com/FootprintAI/selenium-example>

Use Git or checkout with SVN using the web URL.

Open with GitHub Desktop

Download ZIP

# Selenium: Basic Example

```
1 from drivers.webdriver import WebDriver
2
3 from selenium.webdriver.common.by import By
4 from selenium.webdriver.common.keys import Keys
5
6 def do(driver: WebDriver, q: str):
7     driver.get("http://www.python.org")
8     assert "Python" in driver.title
9     elem = driver.find_element(By.NAME, "q")
10    elem.clear()
11    elem.send_keys(q)
12    elem.send_keys(Keys.RETURN)
13    assert "No results found." not in driver.page_source
```

By is for element selection  
Key is for mapping keyboard

# Selenium: Webdriver Inheritance & encapsulation

Type hint

```
1 import os
2 import platform
3 from selenium.webdriver import Chrome, ChromeOptions
4
5 class WebDriver(Chrome):
6     def __init__(self, headless:bool=False):
7         exec_path = None
8         if platform.system() == "Darwin":
9             exec_path = self.__wrap_abspath('chromedriver.darwin')
10            elif platform.system() == "Windows":
11                exec_path = self.__wrap_abspath('chromedriver.exe')
12            elif platform.system() == "Linux":
13                exec_path = self.__wrap_abspath('chromedriver.linux')
14
15        chrome_options = ChromeOptions()
16        # comment this to enable guest browser to show up (for debugging)
17        if headless:
18            chrome_options.add_argument('--headless')
19            chrome_options.add_argument('--no-sandbox')
20
21        Chrome.__init__(self,chrome_options=chrome_options, executable_path=exec_path)
22
23    def __wrap_abspath(self, binary)-> str:
24        return os.path.join(os.path.dirname(os.path.abspath(__file__)), binary)
```

inheritance

adaptor

# Selenium: Cookies & Navigation

- Cookies are small text files recording your identity and some credential information to allow website identify your access.
- If you want to have more privacy, just delete cookies (come along when you closed your browser)

```
1  from drivers.webdriver import WebDriver
2
3  from time import sleep
4
5  def navigation(driver: WebDriver):
6      driver.get("https://www.google.com")
7      sleep(1)
8      driver.get("https://www.python.org")
9      sleep(1)
10     driver.back()
11     sleep(1)
12     driver.forward()
13
14 def get_cookies(driver: WebDriver):
15     driver.get("https://www.google.com")
16     print(driver.get_cookies())
```

# Selenium: Exceptions

- Exception allows your program to know the results are unexpected, so they would get human to be involved.

```
1 from drivers.webdriver import WebDriver
2
3 def exception_handling(driver: WebDriver):
4     from selenium.common.exceptions import TimeoutException, NoSuchElementException
5     try:
6         driver.get("https://www.google.com")
7     except TimeoutException:
8         print("time out")
9     try:
10         driver.find_element_by_id("you-should-not-see-me")
11     except NoSuchElementException:
12         print("no element found")
```

# Selenium: Query methods

Tag

```
<input type="text" name="passwd" id="passwd-id" />
```

Attribute

```
// we have four different ways to locate an individual UI element
```

```
element = driver.find_element(By.ID, "passwd-id")
element = driver.find_element(By.NAME, "passwd")
element = driver.find_element(By.XPATH, "//input[@id='passwd-id']")
element = driver.find_element(By.CSS_SELECTOR, "input#passwd-id")
```

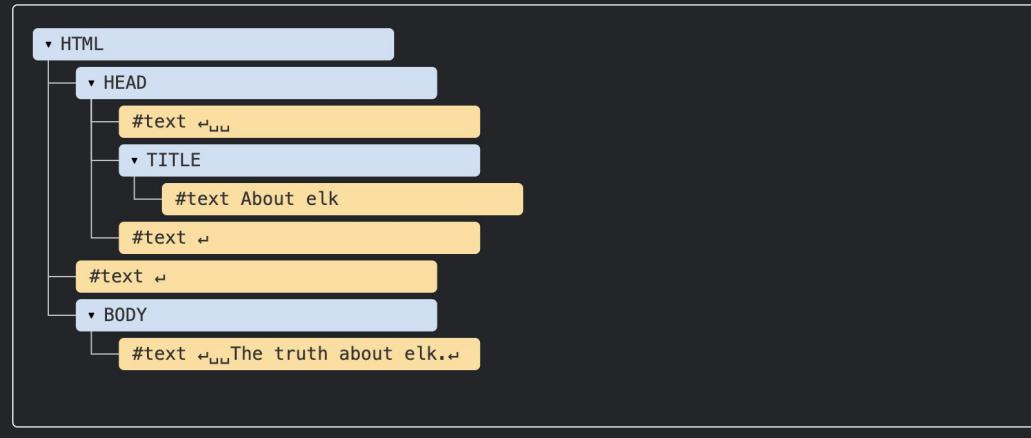
# DOM(Document Object Model) Tree

## An example of the DOM

Let's start with the following simple document:

```
1 <!DOCTYPE HTML>
2 <html>
3 <head>
4   <title>About elk</title>
5 </head>
6 <body>
7   The truth about elk.
8 </body>
9 </html>
```

The DOM represents HTML as a tree structure of tags. Here's how it looks:



Ref: <https://javascript.info/dom-nodes>

# Example: Instagram Login

```
0
7     driver.get("https://www.instagram.com")
8     sleep(5) # wait for the page loading is finished
9     # find username fields
10    username_field=driver.find_element_by_css_selector("input[name='username']")
11    # find password fields
12    password_field=driver.find_element_by_css_selector("input[name='password']")
13    username_field.clear()
14    password_field.clear()
15    # send email in username field
16    username_field.send_keys(email)
17    # send password in password field
18    password_field.send_keys(password)
19    sleep(5)          Always have a random delay to mimic human behavior
20
21    # click login button
22    driver.find_element_by_css_selector("button[type='submit']").click()
23    __post_login(driver)
24
25 def __post_login(driver: WebDriver):
26
```

## Example: Find Stock Price

```
1  from drivers.webdriver import WebDriver
2
3  from selenium.webdriver.common.by import By
4  from selenium.webdriver.support.ui import WebDriverWait
5  from selenium.webdriver.support import expected_conditions as EC
6
7  def do_find(driver: WebDriver, symbol: str):
8      driver.get("https://goodinfo.tw/StockInfo/index.asp")
9      ele = driver.find_element_by_xpath(u"//*[@id='txtStockCode']")
10     ele.clear()
11     ele.send_keys(symbol)
12     driver.find_element_by_xpath(u"//input[@value='股票查詢']").click()
13
14     # the following code will wait until the element is visible
15     try:
16         wait = WebDriverWait(driver, 30)
17         wait.until(EC.visibility_of_element_located((By.XPATH, u"//*[@id='imgKC']")))
18     except:
19         raise Exception("target element is not show up")
```

Or Wait until a certain element has shown up.

# You have a scraping, we have lots of anti-Scraping tools.

- How to bypass Anti-Scraping Tool
  - Keep Rotate Your IP Address
    - By a well-paid proxy service provider
    - By restricting a certain area of your location
  - Use Real User Agent and valid Referrer
    - UA(User-Agent) are http header which is used identify what browser type you used to visit this website.
    - Keep a list of valid UI and use them randomly.
  - Avoid Periodically Requests
    - Keep random intervals between requests, random delay are helpful.
  - Watch Exception and update your code
    - Anti-scraping tool would change DOM structure frequently, making your script failed to find target elements.
- However, Implementing these guidelines could costly if your application is still below a certain scale.

# Shrimping: A data-sharing platform

Shrimping provides a unified way for clients to get human-centric information in a simple, easy, and low cost fashion.

<https://get-shrimping.footprint-ai.com/>

# Web scraping scenarios for Shrimping.

- UI Validation
  - When working with business partners, it is extremely important that your partner has interpreted your product correctly.
- Collecting sell history on an ecommerce platform
  - Track the sell volume of all products on an eCommerce platform.
  - As the number of products could be big (approximately 100M active products), how to get each product's sell records on daily basis is extremely challenging.
- Scraping and analyzing KOL's feeds on social network platforms
  - Find out a KOL and his/her fans preference for retargeting, reselling, or other marketing strategies.
  - Social network platform always implemented anti-bot mechanism, making it hard to collect in a large scale.

# Modeling

Modeling a problem  
with Machine Learning

Modeling is processing your dataset and generating a patterns or distributions to fit your dataset.

# Data Analysis with Panda and Numpy

<https://pandas.pydata.org/docs/>

<https://numpy.org/doc/stable/user/index.html#user>

```
import numpy as np
```

```
import pandas as pd
```

# Pandas Series: One-dimensional array of holding **any** data type

```
>>> import numpy as np
>>> a = pd.Series(np.random.randn(5))
>>> a
0    0.427981
1    0.998562
2    0.649786
3    1.341638
4    0.743008
dtype: float64
>>> a[a > a.median()]
1    0.998562
3    1.341638
dtype: float64
>>> a.size
5
>>> a.array
<PandasArray>
[0.4279808645556574, 0.9985615198039794, 0.6497859868724282,
 1.3416378899984038, 0.7430084840615195]
Length: 5, dtype: float64
>>> a.dtype
dtype('float64')
>>> a.to_numpy()
array([0.42798086, 0.99856152, 0.64978599, 1.34163789, 0.74300848])
>>> a2 = a * 2
>>> a2
0    0.855962
1    1.997123
2    1.299572
3    2.683276
4    1.486017
dtype: float64
>>> a3 = a[:-1] + a2[1:]
>>> a3
0        NaN
1    2.995685
2    1.949358
3    4.024914
4        NaN
dtype: float64
```

# DataFrame: Think like table, work like table...

- Pandas is a Python library and the de-facto standard for working with **structured** tabular data on Python
- Rich Format Supported including CSV, JSON, Parquet, MS EXCEL, ...

asset	max_training	max_validation	mean_training	mean_validation	med_training	med_validation	min_training	min_validation	std_training	std_validation
AAN	2.685913	5.132591	0.046193	0.029751	0.120043	0.237710	-4.324665	-7.955100	1.356045	2.851172
ACC	1.546164	3.473945	0.024481	-0.191317	0.067734	0.000000	-1.493781	-8.192803	0.662752	2.181609
ACGL	1.225804	3.676009	0.035951	-0.350968	0.072673	0.000000	-1.788876	-6.347304	0.607838	1.949640
ADC	1.343051	1.567481	0.118847	-0.090484	0.152312	0.001203	-1.302874	-4.310633	0.603213	1.131439
AEL	3.002969	5.464704	0.026192	-0.094789	0.076017	0.000000	-3.330113	-7.611560	1.290956	2.659290
AFG	1.095037	3.558017	0.005339	-0.387737	0.072613	-0.060772	-2.050200	-7.330084	0.614609	2.013369
AFL	1.159893	3.050265	0.051159	-0.295960	0.072861	-0.038611	-1.414419	-5.840748	0.514807	1.709449
AGM	2.189914	2.572398	0.119186	-0.179405	0.130402	-0.031308	-2.258110	-5.378916	0.944030	1.627170
AGNC	0.934225	2.371371	-0.016914	-0.196734	0.026154	0.159754	-1.373769	-6.580712	0.481505	1.653974

Example of a simple Pandas table with “assets” as index and various numerical columns

# Dataframe Indexing

- **iloc (position-based indexing) vs loc (label-based indexing)**
  - Label-based indexing returns with “match” labels
  - Position-based indexing returns elements with its numeric position index

```
>>> f = pd.DataFrame({'a':[1,2,3,4,5], 'b':[10,20,30,40,50], 'c': [9,8,7,6,5]})  
>>> f.head()  
   a   b   c  
0  1  10   9  
1  2  20   8  
2  3  30   7  
3  4  40   6  
4  5  50   5  
>>> f.loc[2:3, 'b']  
2    30  
3    40  
Name: b, dtype: int64  
>>> f.iloc[2:3, 1]  
2    30  
Name: b, dtype: int64
```

# Dataframe Join Operation

- Join provides a way to intersect / union two Dataframes
  - Use 'On' to specify a specific join key
  - Use 'How' to specify we want to have inner / outer join

```
>>> f = pd.DataFrame({'a':[1,2,3,4,5], 'b':[10,20,30,40,50], 'c': [9,8,7,6,5]})  
>>> g = pd.DataFrame({'e':['I','II','III']})  
>>> f.join(g, how='outer')  
      a    b    c    e  
0   1   10   9    I  
1   2   20   8   II  
2   3   30   7  III  
3   4   40   6  NaN  
4   5   50   5  NaN  
>>> f.join(g, how='inner')  
      a    b    c    e  
0   1   10   9    I  
1   2   20   8   II  
2   3   30   7  III  
>>> f.join(g, how='inner', on="a")  
      a    b    c    e  
0   1   10   9   II  
1   2   20   8  III
```

# Dataframe Multiindex

- Indexing allows you projecting Dataframe, and Multiindex simplify this process

```
>>> adult[adult.columns[:8]].head()
   workclass fnlwgt marital_status occupation relationship race capital_gain capital_loss
education education_num sex      age
10th      6           Female 17          ?  Never-married      ?        Own-child  White     34095      0
          17          Private 17          Never-married Other-service  Own-child  White      0      0
          17          ?          158762 Never-married      ?        Own-child  White      0      0
          17          Private 329783 Never-married    Sales  Other-relative  White      0      0
          17          Private 139183 Never-married    Sales        Own-child  White      0      0
>>> adult.loc[[' Bachelors'], ['label']]
            label
education education_num sex      age
Bachelors 13           Female 19    <=50K
                  19    <=50K
                  20    <=50K
                  21    <=50K
                  21    <=50K
...
                  ...
                  Male   90    <=50K
                  90    >50K
                  90    <=50K
                  90    <=50K
                  90    <=50K
[5355 rows x 1 columns]
>>> adult.loc[[' Bachelors', 13, ' Female', 19], ['label']]
            label
education education_num sex      age
Bachelors 13           Female 19    <=50K
                  19    <=50K
```

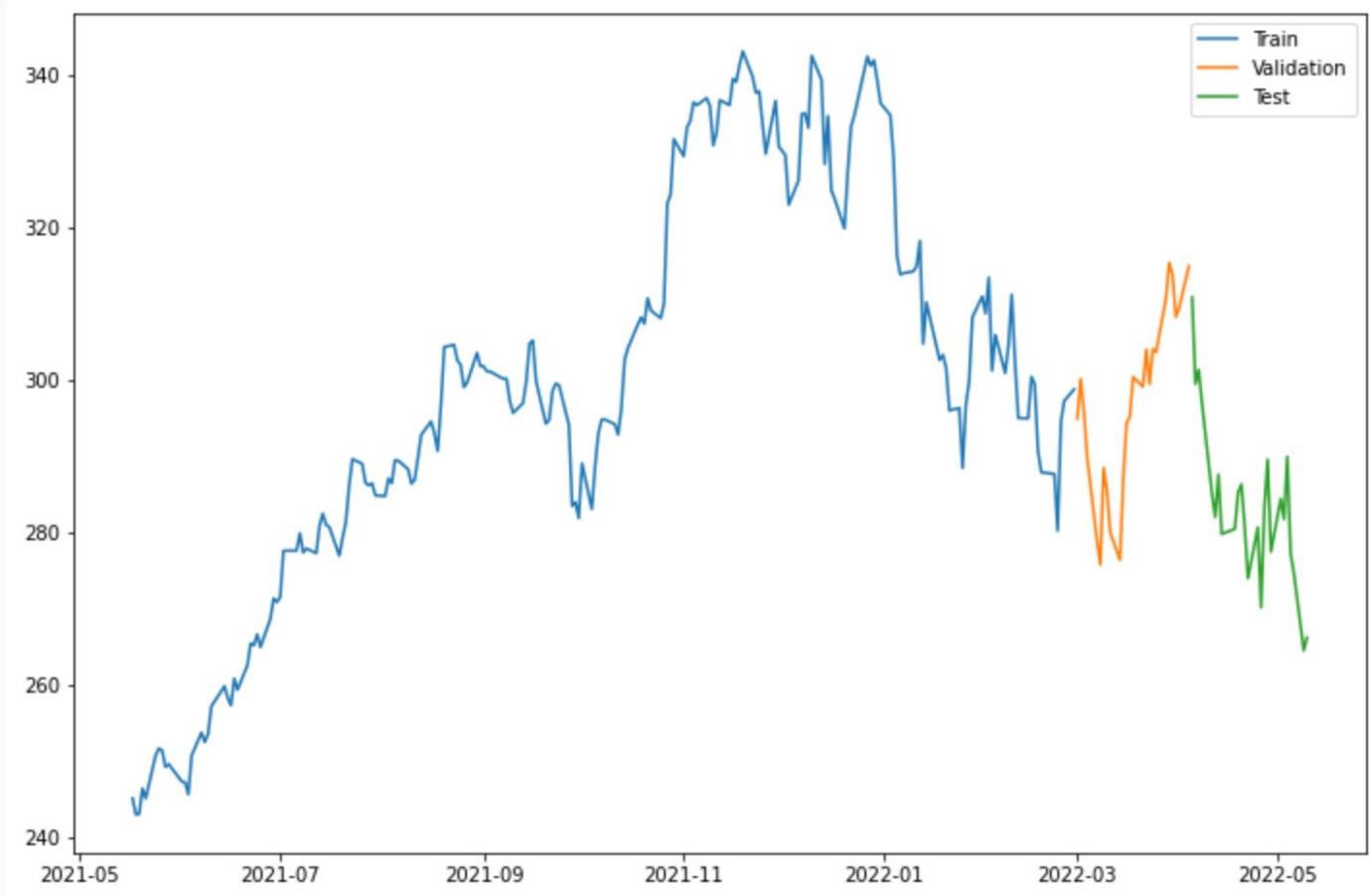
# Dataframe Query

- Query is the most used filter to filter rows by its column's value.

```
>>> adult
   education      education_num  sex  age  workclass  fnlwgt  marital_status  occupation  ...  capital_loss  hours_per_week  native_country  label
0    10th            6        Female  17       ?  304873  Never-married          ?  ...  0  32  United-States  <=50K
1    10th            6        Female  17     Private  169658  Never-married  Other-service  ...  0  21  United-States  <=50K
2    10th            6        Female  17       ?  158762  Never-married          ?  ...  0  20  United-States  <=50K
3    10th            6        Female  17     Private  329783  Never-married          Sales  ...  0  10  United-States  <=50K
4    10th            6        Female  17     Private  139183  Never-married          Sales  ...  0  15  United-States  <=50K
...
5  Some-college  10        Male  84  Self-emp-inc  172907  Married-civ-spouse          Sales  ...  0  35  United-States  >50K
6  Some-college  10        Male  90     Private  52386  Never-married  Other-service  ...  0  35  United-States  <=50K
7  Some-college  10        Male  90     Private  52386  Never-married  Other-service  ...  0  35  United-States  <=50K
8  Some-college  10        Male  90  Self-emp-not-inc  282095  Married-civ-spouse  Farming-fishing  ...  0  40  United-States  <=50K
9  Some-college  10        Male  90       ?  225063  Never-married          ?  ...  0  10  South  <=50K
[32561 rows x 11 columns]
```

```
>>> adult.query("workclass == ' Private' ")
   education      education_num  sex  age  workclass  fnlwgt  marital_status  occupation  ...  capital_loss  hours_per_week  native_country  label
0    10th            6        Female  17     Private  169658  Never-married  Other-service  ...  0  21  United-States  <=50K
1    10th            6        Female  17     Private  329783  Never-married          Sales  ...  0  10  United-States  <=50K
2    10th            6        Female  17     Private  139183  Never-married          Sales  ...  0  15  United-States  <=50K
3    10th            6        Female  17     Private  130125  Never-married  Other-service  ...  0  20  United-States  <=50K
4    10th            6        Female  17     Private  27032  Never-married          Sales  ...  0  12  United-States  <=50K
...
5  Some-college  10        Male  79  Private  121318  Married-civ-spouse  Adm-clerical  ...  0  20  United-States  <=50K
6  Some-college  10        Male  80  Private  87518  Never-married  Prof-specialty  ...  1816  60  United-States  <=50K
7  Some-college  10        Male  81  Private  122651  Married-civ-spouse          Sales  ...  0  15  United-States  <=50K
8  Some-college  10        Male  90  Private  52386  Never-married  Other-service  ...  0  35  United-States  <=50K
9  Some-college  10        Male  90  Private  52386  Never-married  Other-service  ...  0  35  United-States  <=50K
[22696 rows x 11 columns]
```

## How to predict stock price for the next day?



## What sequences means in a series?

Input: [300, 320, 310, 350, 390]

t0	300
t1	320
t2	310
t3	350
t4	390
t5	?

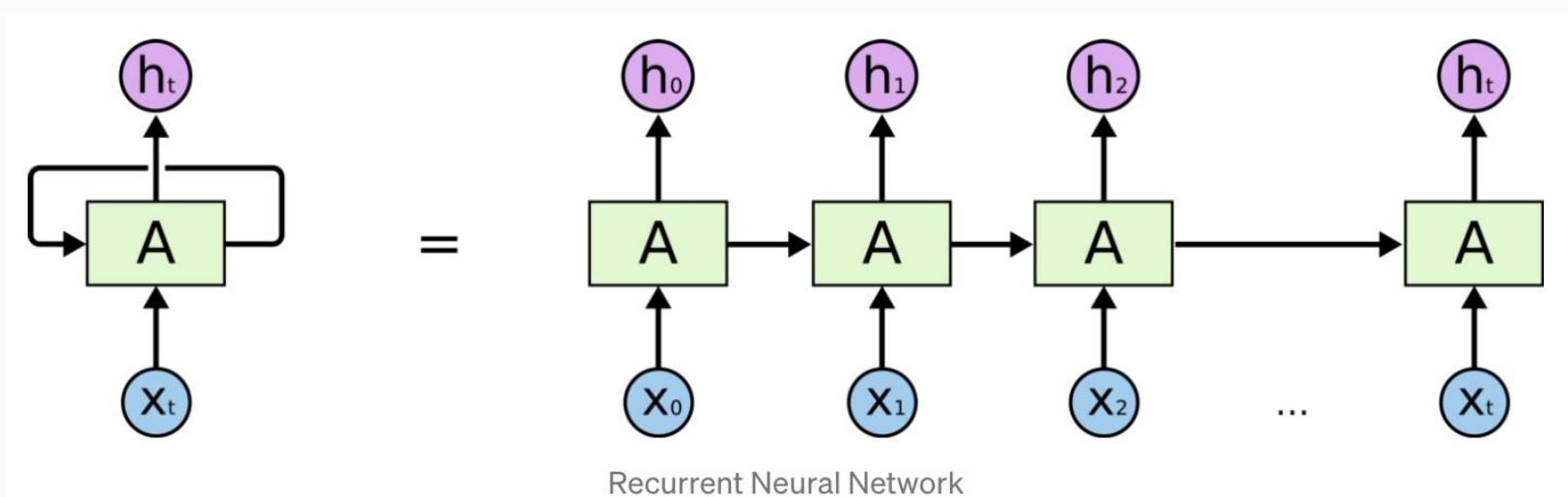
## What sequences means in a sentence?

Input: I am a good guy

t0	I
t1	am
t2	a
t3	good
t4	guy
t5	?

# What Is RNN (Recurrent Neural Network)?

Recurrent Neural Network (RNN) takes decisions on CURRENT ( $X_t$ ) and PREVIOUS ( $X_{t-1}$ ) inputs. Especially useful in topics including machine translation, speech recognition.

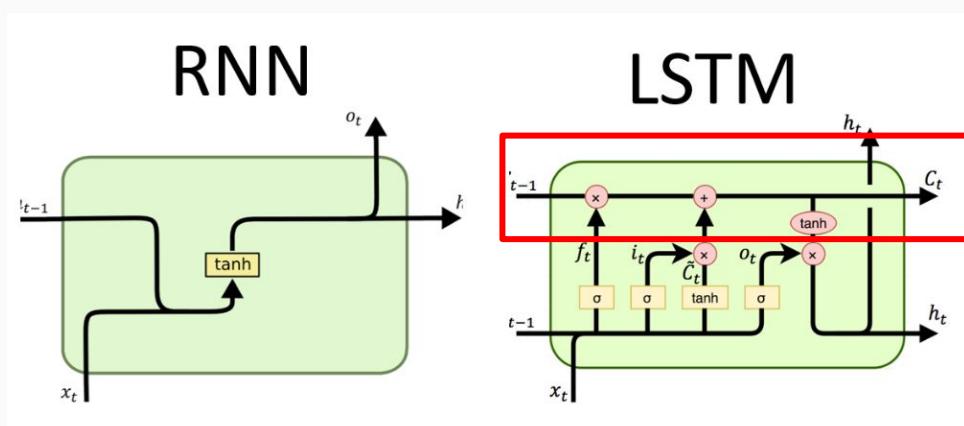


# What Is LSTM (Long-short term memory)?

RNN only remember the latest things from X and it didn't remember(no memory) anything before at the beginning.

LSTM provides an information highway to let the neuron to selectively choose

1. forget from its memory (focus on the current inputs)
2. Listens to what information it added into memory (though information highway)



<https://towardsdatascience.com/introduction-to-recurrent-neural-network-27202c3945f3>

<https://www.quora.com/How-is-LSTM-different-from-RNN-In-a-layman-explanation>

## How to formulate a sequence into a trainable dataset?

t0	t1	t2	t3	t4
300	320	310	350	390

Windows size 3, we looks last three days data

Feature 1	Feature 2	Feature 3	Target
300	320	310	350
320	310	350	390

# Dataframe Visualization

```
plt.figure(figsize=(15, 6))
plt.subplots_adjust(top=1.25, bottom=1.2)

company_list = [MSFT, AAPL, TSLA, VOO]

for i, company in enumerate(company_list, 1):
    plt.subplot(2, 2, i)
    company['Adj Close'].plot()
    plt.ylabel('Adj Close')
    plt.xlabel(None)
    plt.title(f"Closing Price of {tech_list[i - 1]}")

plt.tight_layout()
```



# Code Walkthrough

Short url:

<https://shorturl.at/imRW0>

Original url:

<https://github.com/FootprintAI/kubeflow-workshop/blob/main/tutorials/stockprice-with-lstm/1.visualize-and-build-stockprice-model.ipynb>

What is Tintin?

# Tintin: Machine Learning Platform for Everyone

Tintin is a machine learning platform for easing the gap in building machine learning models with scale.

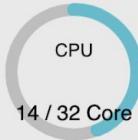
<https://get-tintin.footprint-ai.com/>

## PROJECTS / OVERVIEW

歡迎您!

深度學習訓練開發平台!

### User Resource Overview



### Projects

 New Project



英文語音識別

by [user2@footprint-ai.com](mailto:user2@footprint-ai.com)

最後編輯時間2021/1/26 14:59:31



Cifar-10 影像辨識

by [user2@footprint-ai.com](mailto:user2@footprint-ai.com)

最後編輯時間2021/1/18 15:52:02



# Tintin: Machine learning platform for everyone (2/3)



三 深度學習訓練開發平台

PROJECTS / Cifar-10 影像辨識 / JOBS

Search...



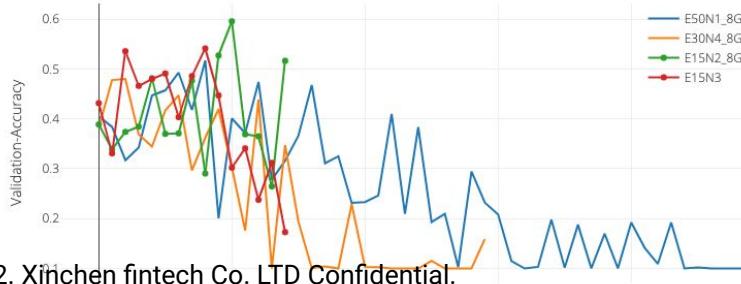
4 selected

Name	Job Status	Onlined	Duration	ScheduleAt	StartAt	EndAt	Metrics
E50N1_8GB	✓	🕒	1:43:30	-	2021/1/13 24:36:39	2021/1/13 02:20:09	▼
E30N4_8GB	✓	🕒	0:43:36	-	2021/1/13 24:26:46	2021/1/13 01:10:22	▼
E15N2_8GB	✓	🕒	0:43:23	-	2021/1/12 23:03:29	2021/1/12 23:46:52	▼
E15N3	✓	🕒	0:26:24	-	2021/1/12 22:56:24	2021/1/12 23:22:48	▼

Rows per page: 5 ▾ 1-4 of 4 < >

Comparison

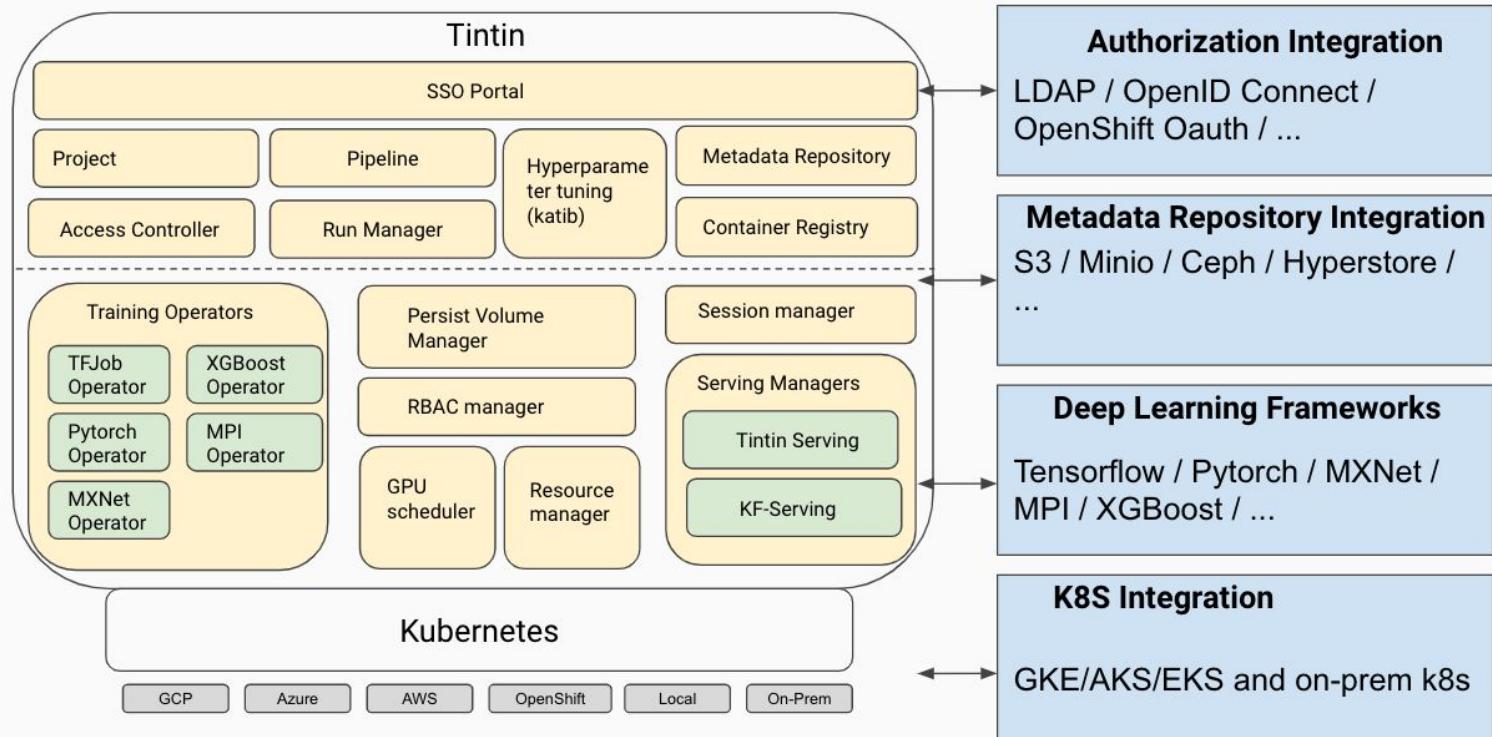
Validation-Accuracy Comparison



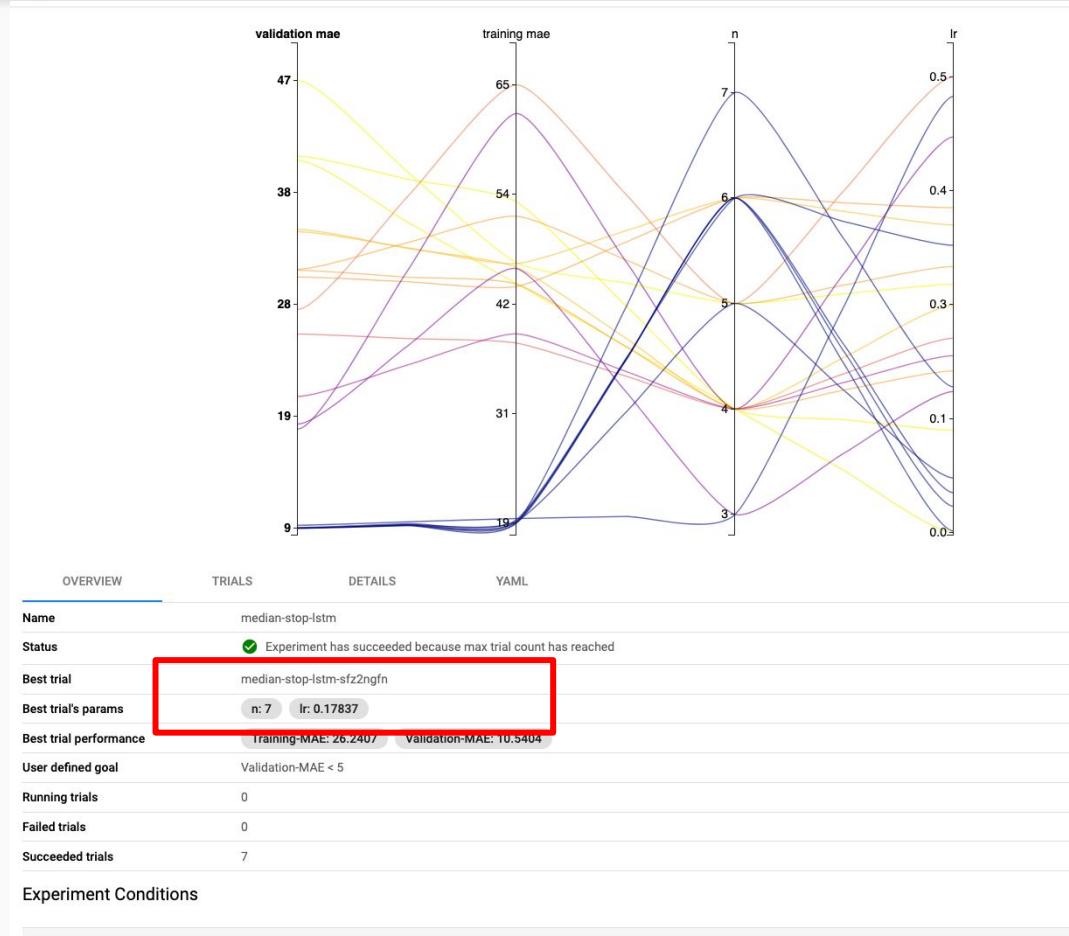
Training-Accuracy Comparison



Empower EVERYONE to build and deploy model easily with Kubeflow and Kubernetes



# Parameter Optimization with Katib



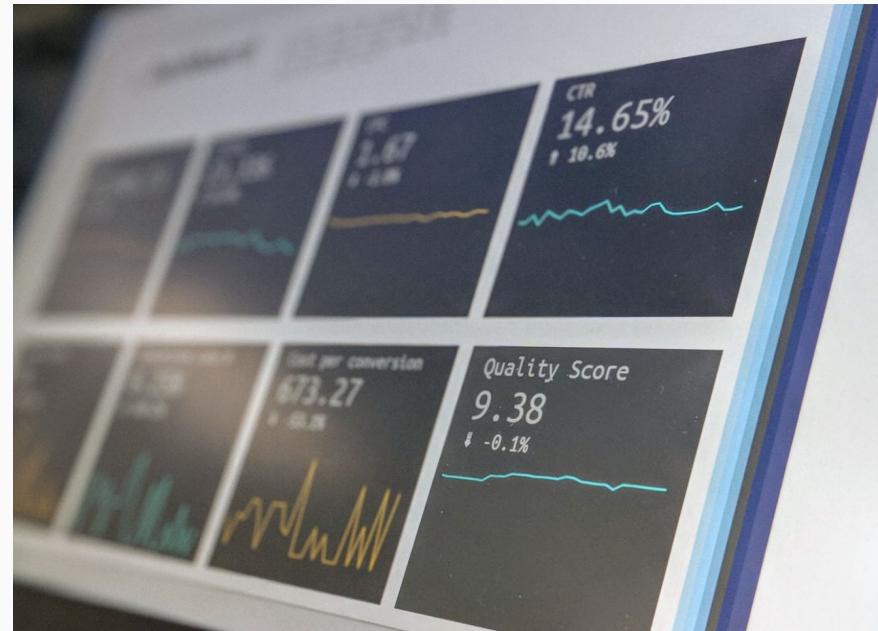
# Serving

## Serving machine learning models with low cost

Serving is the last mile to publish your machine learning models to the public.

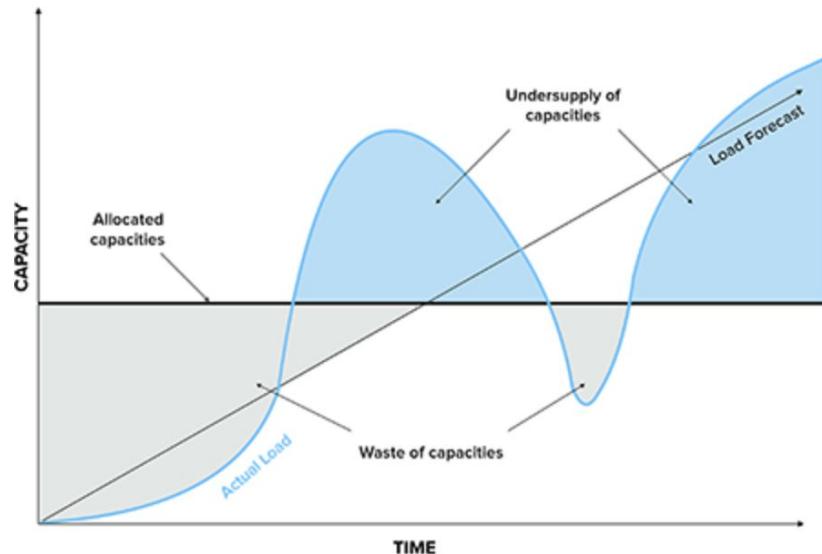
# Model Serving

- How hard it could be to serve ML models in production scale?
  - Scale vs Cost
  - Seamless Rollout
  - Canary Rollouts
  - Service/Model monitoring



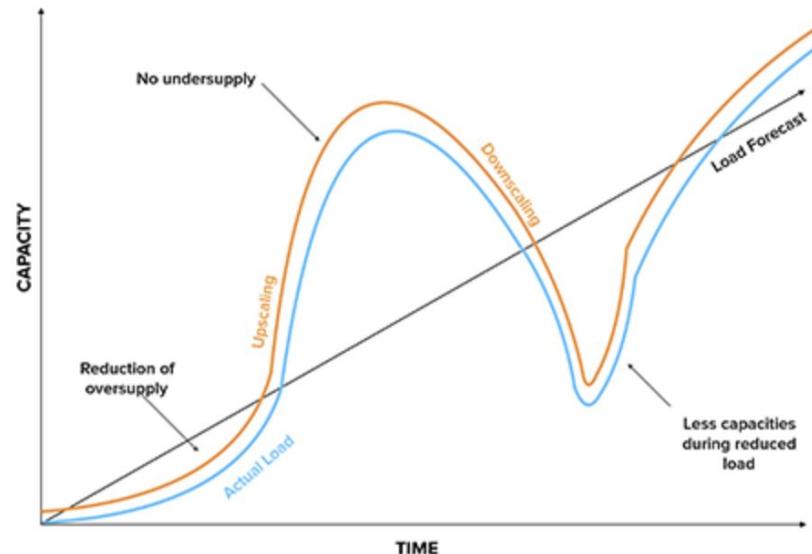
# Cost, Cost, and Cost

Static Architecture



Static architectures are based on estimated load expectancy and are not flexible enough to adapt to unexpected load peaks or lulls.

Auto Scaling Architecture



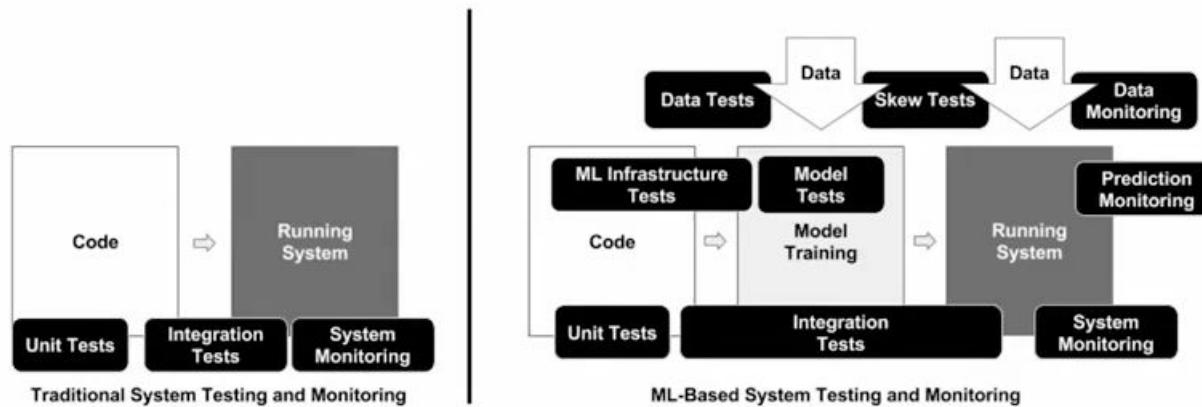
Auto Scaling is the most cost-efficient solution for a fluctuating load. High performance and thereby user satisfaction are retained at all times.

# Why machine learning on Kubernetes?

- Composability
  - Each stage are independent systems and are able to compose together
- Portability
  - Dev/Staging/Prod
  - Laptop/Edge/Cloud environment
- Scalability
  - Hyperparameter tuning, production workloads

# How Involving Machine Learning model could change the current software design?

## Traditional vs. ML infused systems



ML introduces two new assets into the software development lifecycle – **data** and **models**.

**Oh, you want to use ML on K8s?**

**Before that, can you become an expert in:**

- Containers
- Packaging
- Kubernetes service endpoints
- Persistent volumes
- Scaling
- Immutable deployments
- GPUs, Drivers & the GPL
- Cloud APIs
- DevOps
- ...

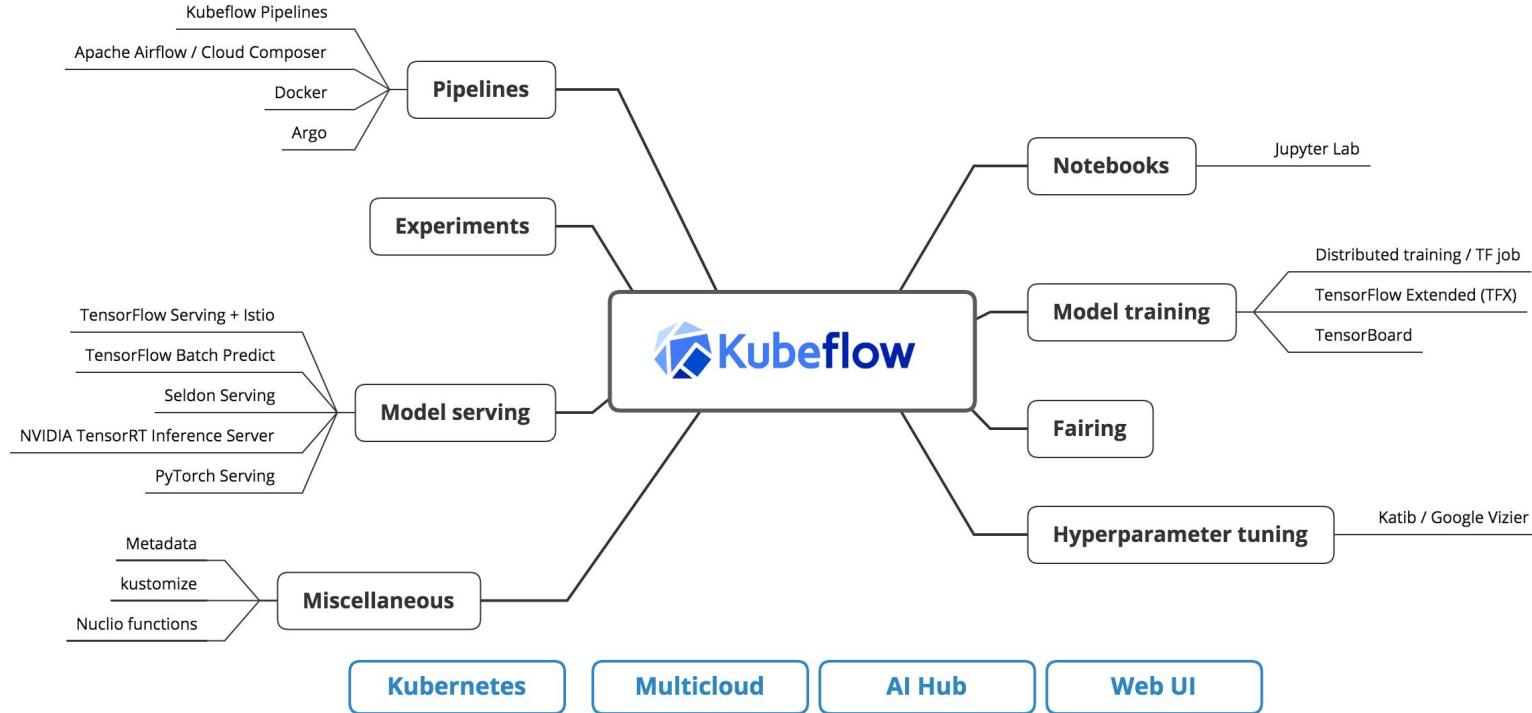




Kubernetes + ML  
= Kubeflow

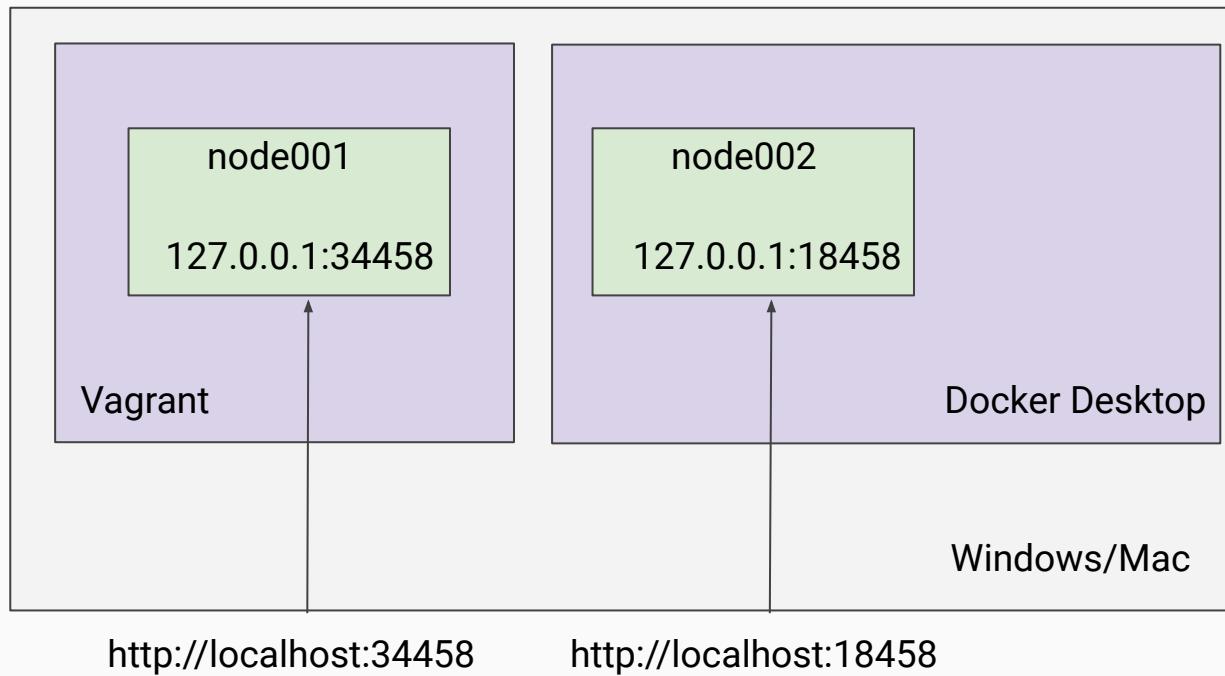
The Kubeflow project  
is dedicated to making  
deployments of  
machine learning (ML)  
workflows on  
Kubernetes simple,  
portable and scalable.

# Architectures



# Multikf: One-click Installation for provisioning multiple kubeflow clusters

- Multikf: <https://github.com/footprintai/multikf>



# Multikf: One-click Installation

```
// Prerequisite: install dockerd (windows)
https://github.com/FootprintAI/kubeflow-workshop/blob/main/install/windows/dockerd.bat.md

// for windows user
// install multikf
wget https://github.com/FootprintAI/multikf/raw/main/build/multikf.windows.exe
chmod +x multikf.windows.exe

// add an instances
./multikf.windows.exe add node002

// connect kubeflow
./multikf.windows.exe connect kubeflow node002
>> now you can open http://localhost:38541
```

# Code Walkthrough

Short url:

<https://shorturl.at/jAV59>

Original url:

[https://github.com/FootprintAI/kubeflow-workshop/tree/main/tutorials/ks  
erve](https://github.com/FootprintAI/kubeflow-workshop/tree/main/tutorials/ks<br/>erve)

What is KaFeiDo?

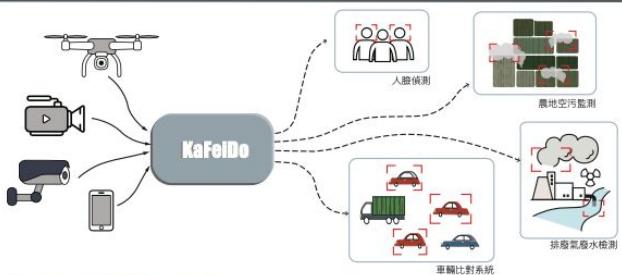
# KaFeiDo: A Machine Learning Platform Built for Green Economy



KaFeiDo is a machine learning platform aiming at saving costs on hardware and energy while providing automation for machine learning models.

## KaFeiDo：智慧節能即時多模型同步推論引擎

我們的一站式平台讓客戶可以選定既有的模型或自行上傳預先訓練的模型進行即時部署。兼具無運算架構與獨家微模型服務節能部署方案，不僅能替客戶省下在模型維運上的人事費用，更可以省下多餘的硬體與電力成本。



### Features highlighted

#### POINT 01

#### KaFeiDo 是您的模型部署好夥伴

KaFeiDo提供模型部署流水線模版將選定的模型數秒之內將模型服務化。由於模型服務化後與後續維護議題是軟體工程的問題而非資料科學家的專業範疇，KaFeiDo將整個過程專業且簡化，透過最佳的實作實例，讓資料科學家可以更專注在模型開發上，以提升客戶的核心價值。

#### POINT 02

#### 異質性多模型同步即時推論

KaFeiDo支援 Triton/Tensorflow/Pytorch 等主流推論框架，讓資料科學家可以使用自己熟悉的框架進行模型開發與部署，讓開發環境不再成為阻礙。

#### POINT 03

#### 無運算架構與水平擴展優勢

KaFeiDo提供無運算架構(Serverless architecture)與水平擴展模組，讓模型服務化不僅能以更低的成本運行，並隨時依據尖峰請求而提高服務量能。由於商品化機器學習模型的趨勢到來，模型數量增長的速度會遠大於硬體增長的速度，KaFeiDo的自動化與需求導向的硬體資源與模型管理機制能讓模型服務依據其請求量能提供適當的計算資源，並在模型服務閒置時將計算資源最小化。

#### POINT 04

#### 微模型服務節能部署架構

KaFeiDo獨家技術微模型服務架構(Micro-model architecture)，有別於傳統的集成式模型推論架構(Monolithic architecture)，微模型服務架構更能減少所需的硬體規格，還能有效降低電力成本。

## KaFeiDo如何運作

適用場景: 社區/學校/醫院/商場/工廠/企業場所  
部署方案: 落地部署(On-prem) / 藝端服務(SaaS)



## KaFeiDo客戶案例

### 永續智慧城市監控中心

藉由匯集多個資料流與多種偵測模型進入即時推論框架，KaFeiDo依據其推論結果觸發警事件通知相關人員，來達成隨時(24/7/365)隨地(簡訊/郵件通知)的分散式監控模式。隨著永續環境概念意識逐漸抬頭，環境監控(如空汙監控、工廠排放廢氣等)更顯得其重要，而如何將智慧監控導入智慧城市變得是一個極嚴峻的問題。傳統的監控中心主要將各個(如攝影機)資訊匯集至單一控制台以便保全人員監控以及當事件發生時提供適當的協助，但隨著監控範圍逐漸拉大(如從閉路電視攝影機到無人機拍攝，從單點監控到場域監控等)，長期依賴保全人員的監控方式除了日益劇增的人力成本以外，也無法長期維持高標準監控。



專業的軟體架構與智慧的節能方案讓您的事業在導入AI上不僅容易且更負擔得起!

## 聯絡我們

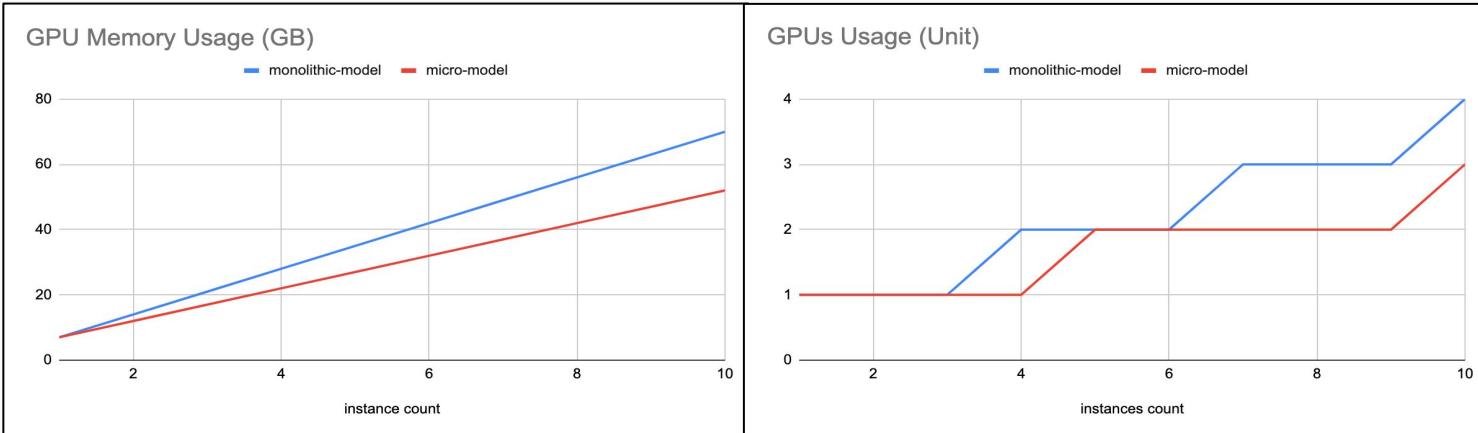
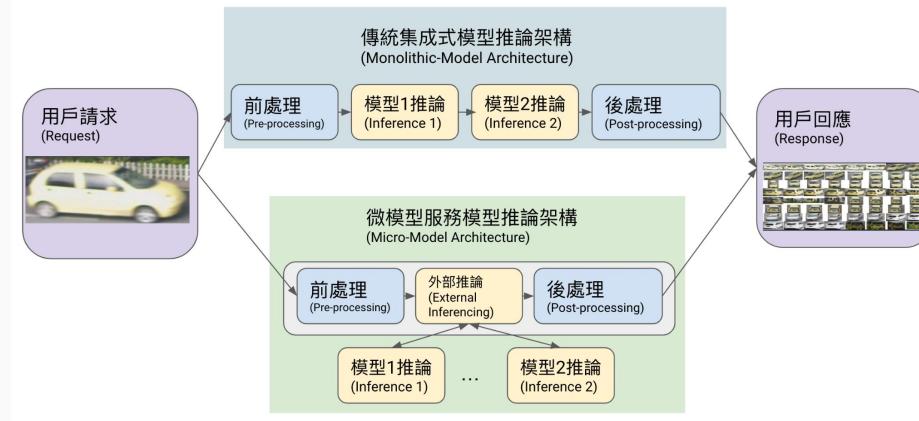
本公司專營機器與深度學習平台 / 網路資料中台 / 多資料模型推論架構 / 客製化模型等服務。  
地址 : 103台北市大同區承德路三段287-2號 email : kafeido@footprint-ai.com

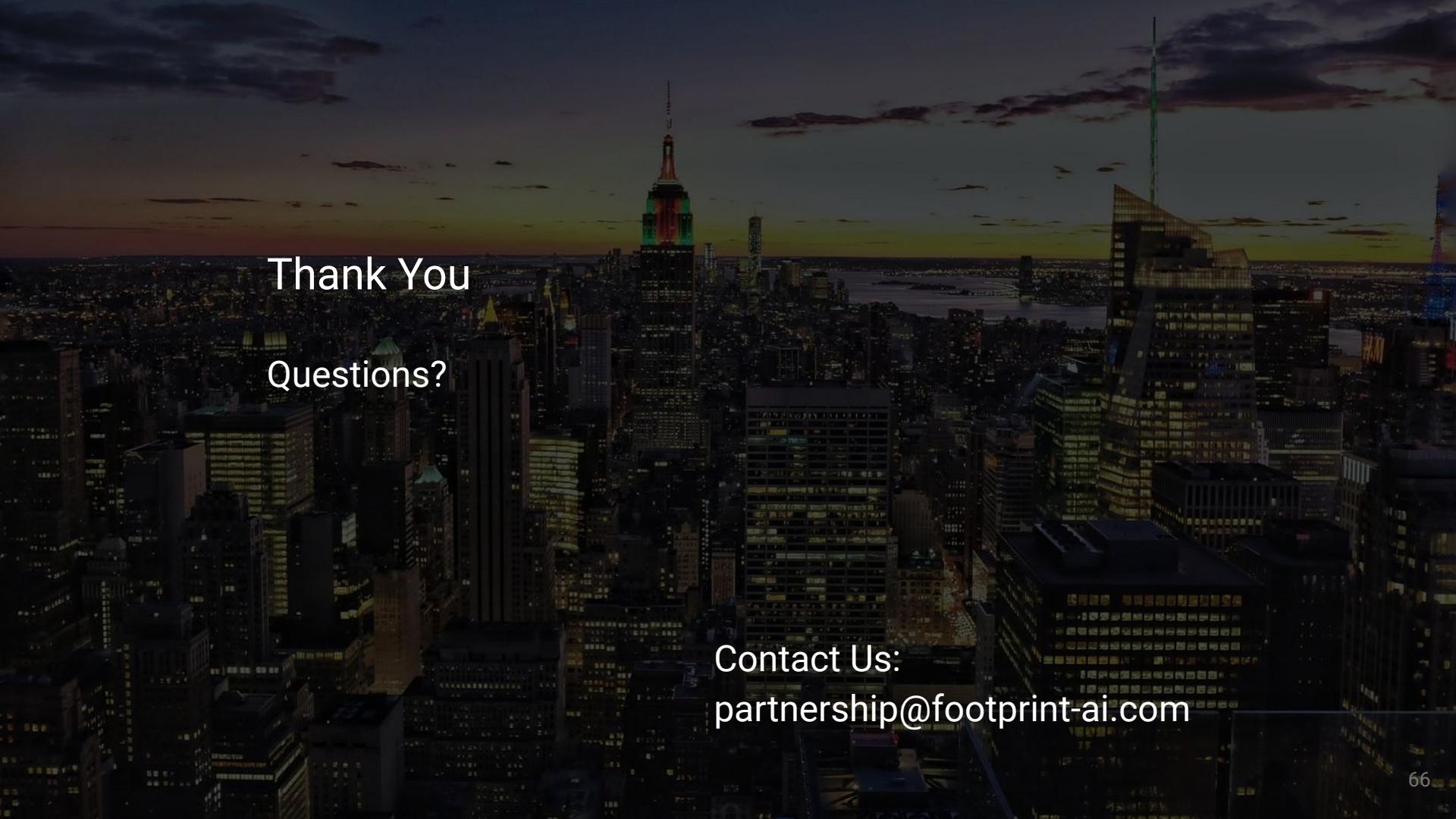


Tintin  
Machine Learning Platform for Everyone



# How we save cost with micro-model architecture?



The background of the slide is a photograph of a city skyline at night, likely New York City, with the Empire State Building prominently visible. The sky is a mix of dark blues and purples with some warm orange and yellow hues from the setting sun.

Thank You

Questions?

Contact Us:  
[partnership@footprint-ai.com](mailto:partnership@footprint-ai.com)



***“The Best Engineers  
Are Lazy”***

-Ancient Engineering Proverb

# Materials

- Slides:
  - <https://github.com/FootprintAI/talks/tree/main/slides>
- Multikf
  - <https://github.com/FootprintAI/multikf>
- Kubeflow Workshop
  - <https://github.com/footprintai/kubeflow-workshop>