



Machine Learning Workshop

Dr. Vinci Chow, CUHK

Teaching Staff

- Instructor: Vinci Chow
 - Office: ELB 1015
 - Office hours: <https://www.ticoneva.com/booked>
 - Email: vincichow@cuhk.edu.hk

Objectives

- Learn Python as a programming language for data analysis
- Get familiar with basic machine learning ideas and tools
- Able to conduct data analysis with real data

Objectives

- Learn Python as a programming language for data analysis
- Get familiar with basic machine learning ideas and tools
- Able to conduct data analysis with real data

Big Data

Machine
Learning

AI

Data Analytics

What's in your mind when
you hear the term "A.I.?"





We've Suggested Tags for Your Photos

We've automatically grouped together similar pictures and suggested the names of friends who might appear in them. This lets you quickly label your photos and notify friends who are in this album.

Tag Your Friends

This will quickly label your photos and notify the friends you tag. Learn more



Who is this?

Who is this?

Who is this?



Who is this?

Who is this?

Who is this?

re

Taco Tuesday

o

Jacqueline Bruzek ×

A

Taco Tuesday

t

Hey Jacqueline,

a

Haven't seen you in a while and I hope you're doing well.

at



What Powers Data Science?



Statistics



**Machine
Learning**



Big Data

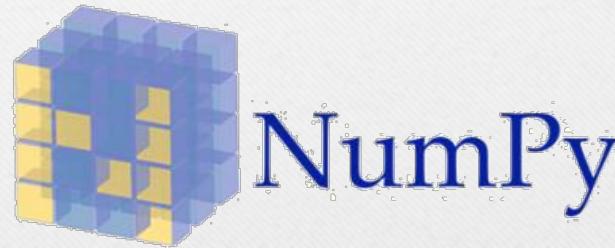
What Power Data Science?



Estimating
Unknown Effects

Statistics

What Power Data Science?



Statistics



StatsModels
Statistics in Python

What Power Data Science?

Make
Accurate
Predictions



**Machine
Learning**

Research on
how to do so
better, faster and
with less data

What Powers Data Science?



Machine
Learning



ONNX



TensorFlow

What Power Data Science?

How big is
Big Data?



Big Data

Age of Big Data

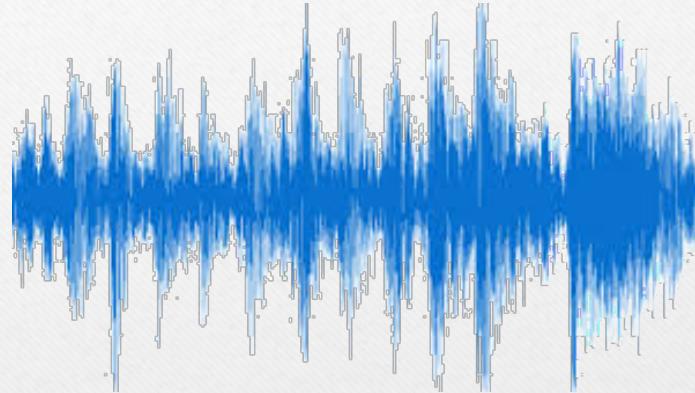
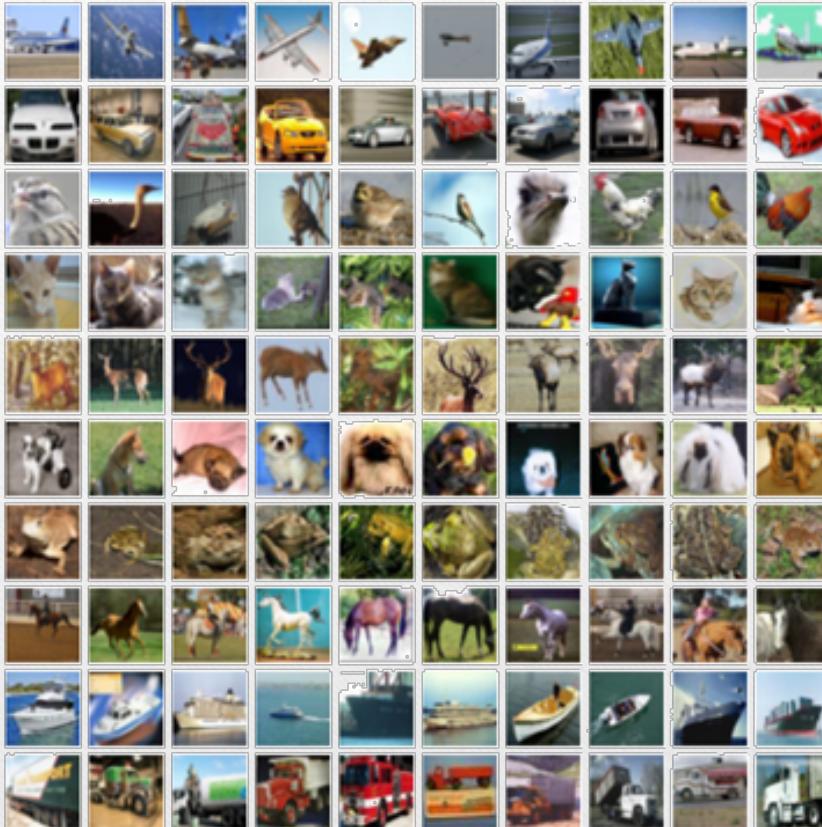


Image Recognition

- **ImageNet**
14 million images
- **CIFAR-10**
60,000 images
- **COCO** (Microsoft)
330,000 images
- **Open Images Dataset**
(Google)
9M images



<http://www.image-net.org>

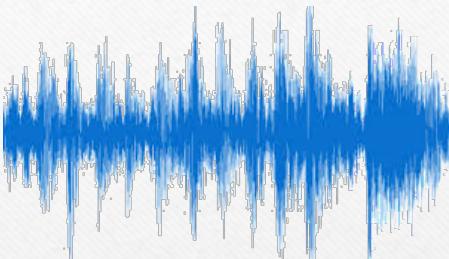
[https://www.cs.toronto.e
du/~kriz/cifar.html](https://www.cs.toronto.edu/~kriz/cifar.html)



<http://cocodataset.org>

[https://storage.googleapis
.com/openimages/](https://storage.googleapis.com/openimages/)

Transcribing Speech to Text

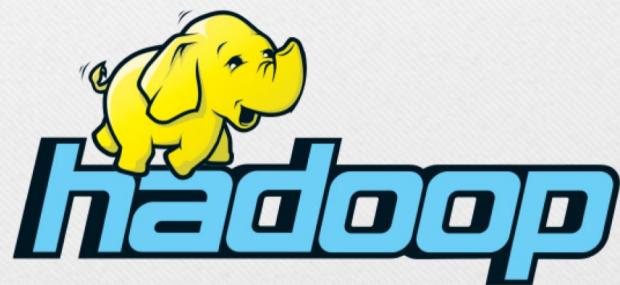


“Call the salon”

- Deep Speech Engine from **Baidu** 百度
- Trained on **11,940 hours** of speech
- Takes **3-6 weeks** to train on a single computer

Source: “Deep Speech 2: End-to-End Speech Recognition in English and Mandarin”

What Power Data Science?



Big Data

How to Take Advantage of AI?

What is AI Good at?

Today's AI systems are only good at specific task. We are very far from **Artificial General Intelligence**.

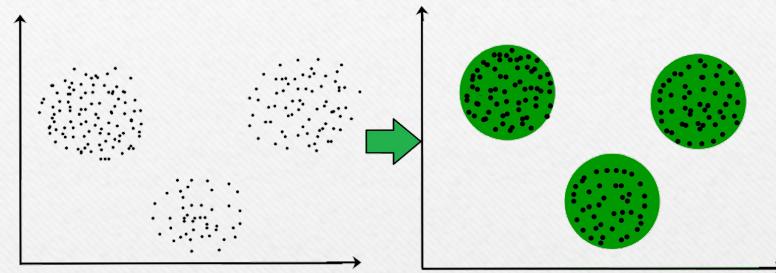


What is AI Good at?

- AI is generally very good at **making predictions**.
 - What is contained in this image?
 - What does this line of Chinese mean in English?
 - What is correct response to the received inquiry?
 - What is the chance of this person having lung cancer?
 - What is the best move in this game of Go?
 - What is the shape of this protein sequence?
 - What is the solution of this differential equation?

What is AI Good at?

- AI is getting better at **generating content.**
 - Automatically categorizing your data.
 - Generating and modifying images to a style of your choice.
 - Simple writing such as short news articles.



O1
生活



Google AI

學啲喇波仔!
(設計對白)



李嘉誠基金會

3星期批10億靠 Google AI 幫手!



Chatbot 互動查詢





NEWSLETTER

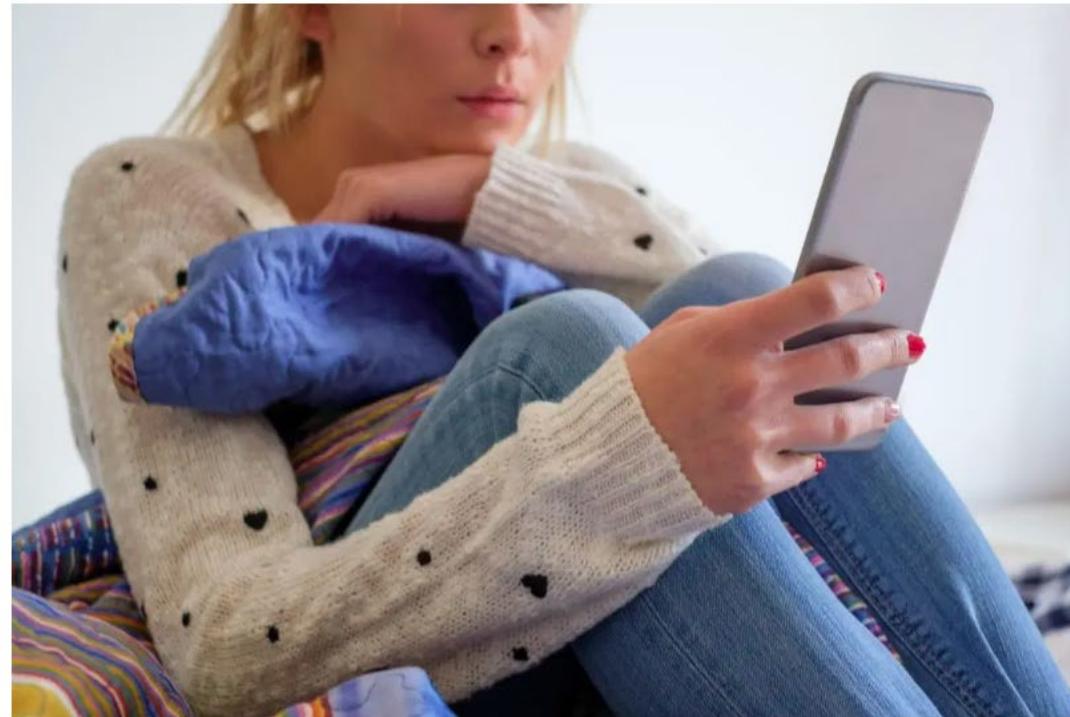
MEDICALNEWS TODAY



Using Facebook to predict depression

Written by [Ana Sandoiu](#) on October 17, 2018 - Fact checked by Paula Field

New research uses over half a million Facebook status updates to predict depression diagnoses in people at risk.



Your Facebook posts may predict whether you will develop depression.

24 SEP 2019

The potentials of...



Adam Booth

Data scientist, AI

AI has a bit of an image problem. It's got a bit of a bias to Brexit, it feels like there's been a bit of a backlash against it. So, when I read...

Improvement in forecast accuracy



Demand

Take account of 75% and make in one forecast



Watch

AI turns you

Using machine learning and Symphony RetailAI's demand forecasting improves forecast accuracy. It allows you to optimize planning for inventory and spoilage, as well as patterns and anomalies.

3x

Demand Forecasting and Planning

Leverage Artificial Intelligence for dramatic improvement in forecast accuracy and planner adoption

Consumer Goods organizations struggle with using multiple versions of the demand signal, coming from multiple systems, to drive their decisions resulting in margin leakage and service level issues.

Antuit has the industry's first driver-based Demand Forecasting and Planning cloud solution, powered by Artificial Intelligence and Machine Learning. We deliver dramatic forecast accuracy improvement and enable planner adoption. When combined with DSD Predictive Ordering and Trade Promotion, the entire solution delivers the intelligence to predict and shape demand.

40% to 60%

No Touch Planning

↑ 2500 BPS

Daily Forecast Accuracy

↑ \$60M

Gross Margin

What is AI **NOT** Good at?

- AI is generally not so good at **explaining things**, partly because complex models are hard to understand.
- If you are interested in pinpointing the relationship between two variables, traditional statistical analysis would probably be more suitable.

What is Your Objective?

- Are you trying to:
 - Let the data tell you what relationship is contained within it?
 - **Unsupervised learning** (not our focus)
 - Trying to predict a continuous variable?
 - Supervised learning - **regression**
 - Trying to predict a discrete variable?
 - Supervised learning - **classification**

SCRP

HPC Cluster

- Run your code on the cloud
- All necessary software installed and maintained
- GPU access
- scrp.econ.cuhk.edu.hk

