

# Introduction to Code-Free Machine Learning

Good Morning!

- 1) Download the presentation slides and activity worksheets at <https://bit.ly/3m8XDgb>
  - 2) If your nickname used in Teams is different from the registered name, please drop a message on the channel with the registered name for attendance tracking purpose.
  - 3) We will start at 9am sharp
- Sit back and relax for now 😊



# Download from Github

zacktohsh / AICFML\_AUG2021 Private

Code Issues Pull requests Actions Projects Security Insights Settings

main ▾

1 branch

0 tags

Go to file

Add file ▾

Code ▾

zacktohsh Main

a35287c 2 minutes ago 3 commits

Additional_Resources	Main	4 minutes ago
.gitattributes	Initial commit	11 minutes ago
AICFML_Activity.v0.3.3.pdf	Main	4 minutes ago
AICFML_Presentation_v0.3.3.pdf	Main	4 minutes ago
Automobile price data _Raw_.csv	Main	4 minutes ago
Car damage dataset.zip	Main	4 minutes ago
README.md	Main	2 minutes ago

README.md



An Introduction to Code-Free Machine Learning (August 2021)



# Programme

---

Section 1:	What is Machine Learning Machine Learning Workflow
Section 2:	Activity 1 – First Machine Learning with Azure
Section 3:	Activity 2 - Deploying your experiment as a Web Service & Make Prediction using Excel
	Lunch Break
Section 4:	Transfer Learning Computer Vision: Activity 3 – Car Damage Assessment Classification
Section 5:	Natural Language Processing Activity 4 – Creating a Sentiment Analyser
Section 6:	Linking them together
Section 7:	Debrief



# Introduction of trainer

---



**Name**  
Zack Toh

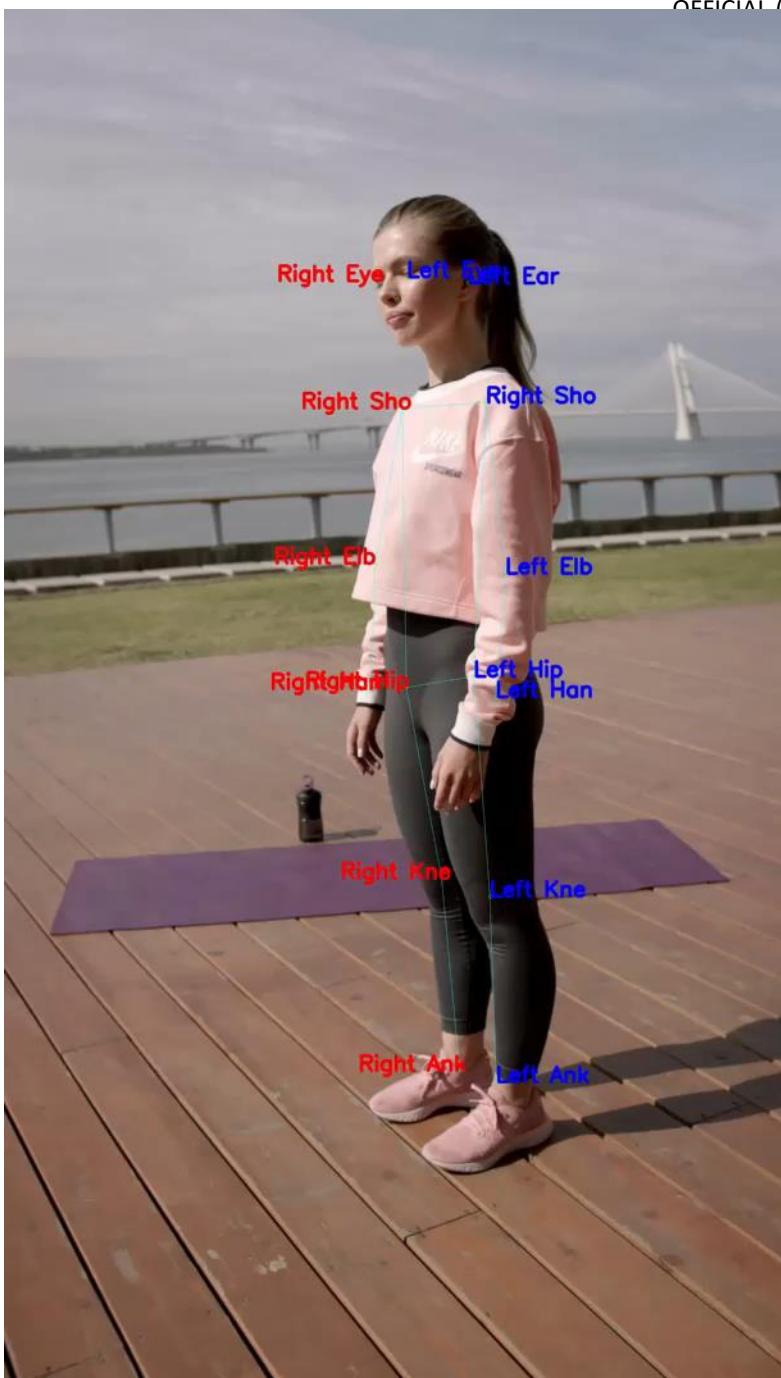
**Telegram**  
[@zacktohsh](https://t.me/zacktohsh)

**Email**  
[Zack\\_toh@rp.edu.sg](mailto:Zack_toh@rp.edu.sg)



# Past Projects (Crowd Detection) (18RIGO09)

**SSDv1****SSDv2****Yolo**



## Pose Recognition

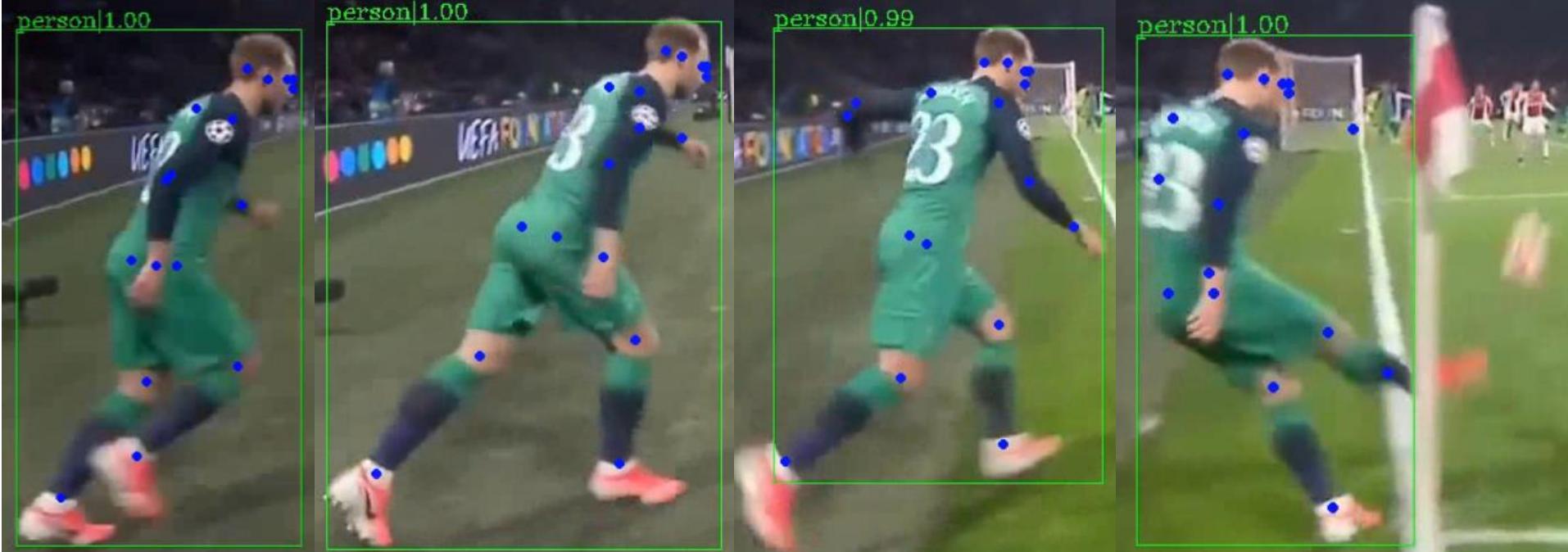
---

Video obtained from: <https://www.pexels.com/>

Pose generated with Posenet

# Deep Learning for Sports Tagging (19RBF07)

Contextual awareness is the holy grail of computer vision. It aims at equipping a machine with the ability of deciphering what is happening in any given image. The automated tagging in sports is a new emerging area in sports industry. This project would contribute tremendously to the productivity and efficiency of sports tagging and dramatically reduce the labor needed in curating the sports statistic. The proposed prototype has high commercialization potential as shared by current sports tagging companies (e.g., Dartfish SA, Hudl and Prozone).



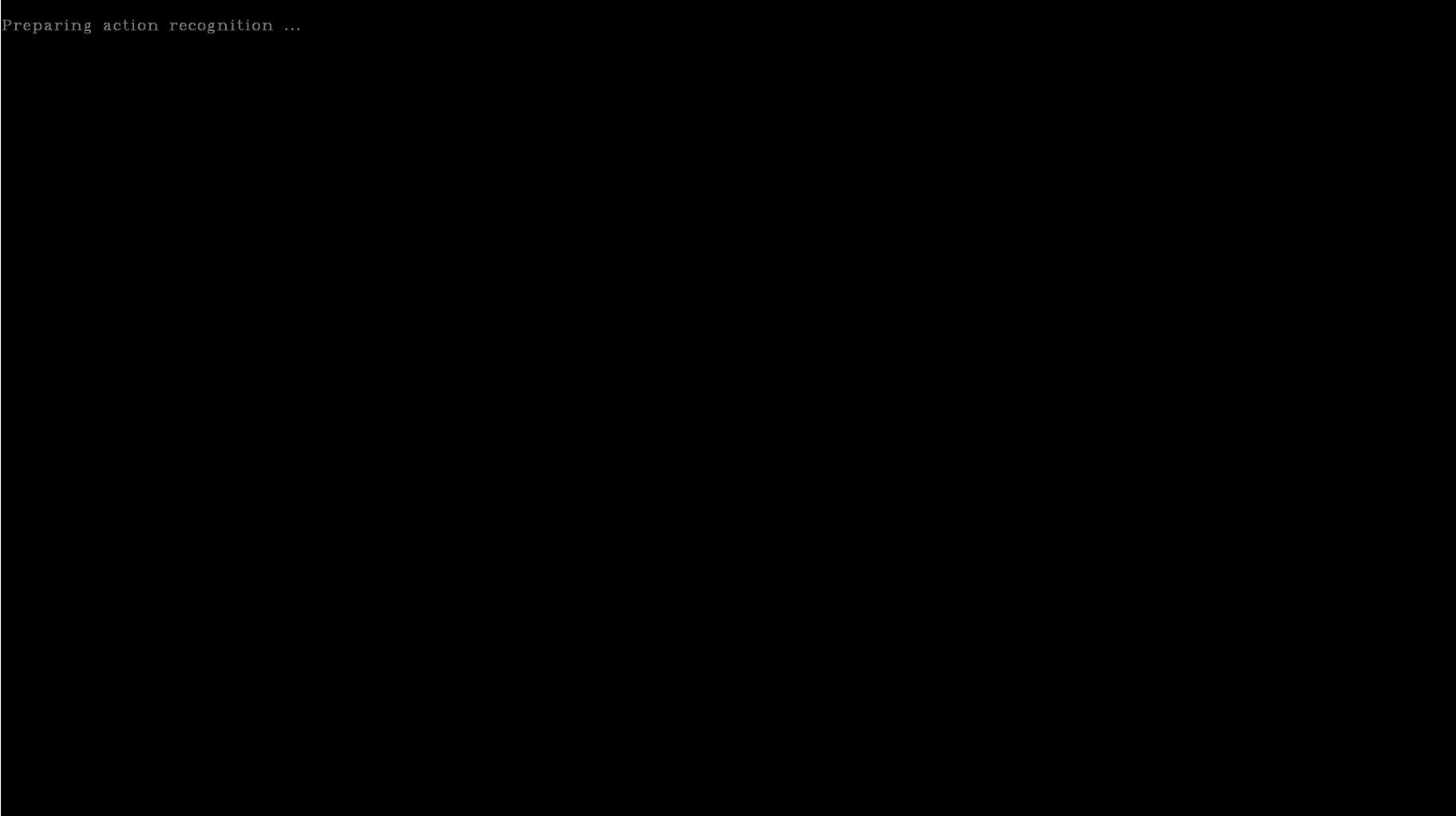
# Processed Throw-in



# Outcome of Project

---

Preparing action recognition ...





# Innovative Approach and Value (PA13)





# Testing of Model Created

---



# Talk to flower

---



[https://www.youtube.com/watch?v=nsPQvZm\\_rgM](https://www.youtube.com/watch?v=nsPQvZm_rgM)

# Plants can 'communicate' with humans, move, and exhibit distress through electrical signals



SINGAPORE - Imagine a day when crops are able to tell you they are thirsty or when you can instruct a plant to delicately pick up an item.

Far-fetched? Scientists in Singapore have actually created a way for humans and plants to communicate with each other - via a smartphone, no less.

The team at Nanyang Technological University (NTU) did this by developing a small conductive material that allows electrical signals to enter and leave the plant.

**THE STRAITS TIMES**

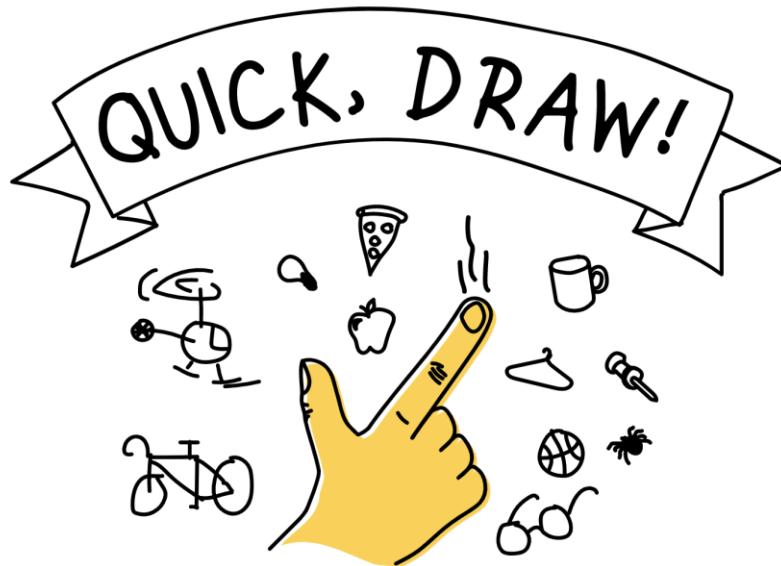
PUBLISHED MAR 16, 2021, 2:37 PM SGT

As part of an experiment, electrical signals were transmitted to the plant using a smartphone, causing the Venus flytrap to close its leaves on demand. PHOTO: NTU SINGAPORE



# Quickdraw Game

<https://quickdraw.withgoogle.com>



Can a neural network learn to recognize doodling?

Help teach it by adding your drawings to the [world's largest doodling data set](#), shared publicly to help with machine learning research.

Let's Draw!

Optional Activity

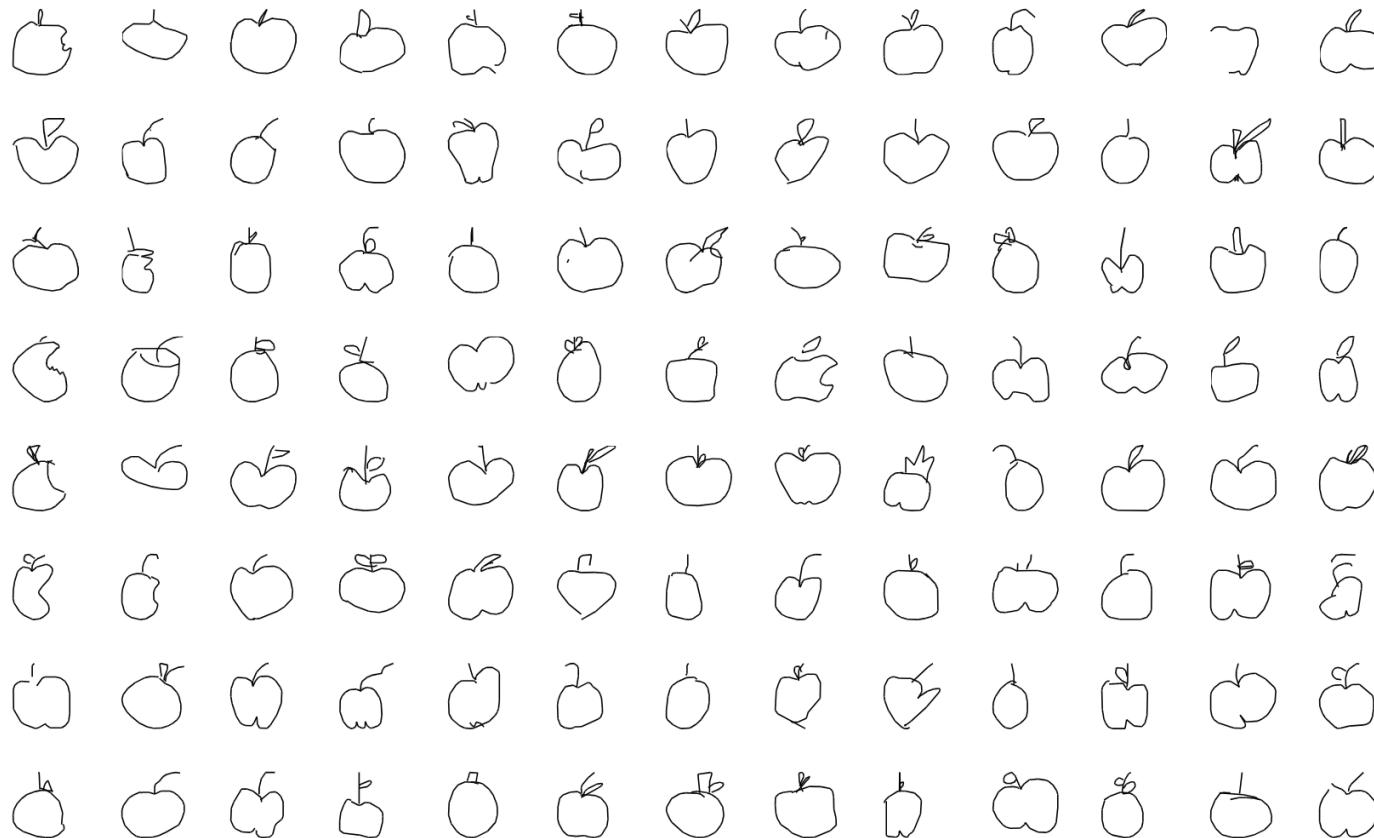




# How does ML work in QuickDraw?

---

- <https://quickdraw.withgoogle.com/data/apple>





# Bias Bias Bias

---

## When It Comes to Gorillas, Google Photos Remains Blind

Google promised a fix after its photo-categorization software labeled black people as gorillas in 2015. More than two years later, it hasn't found one.



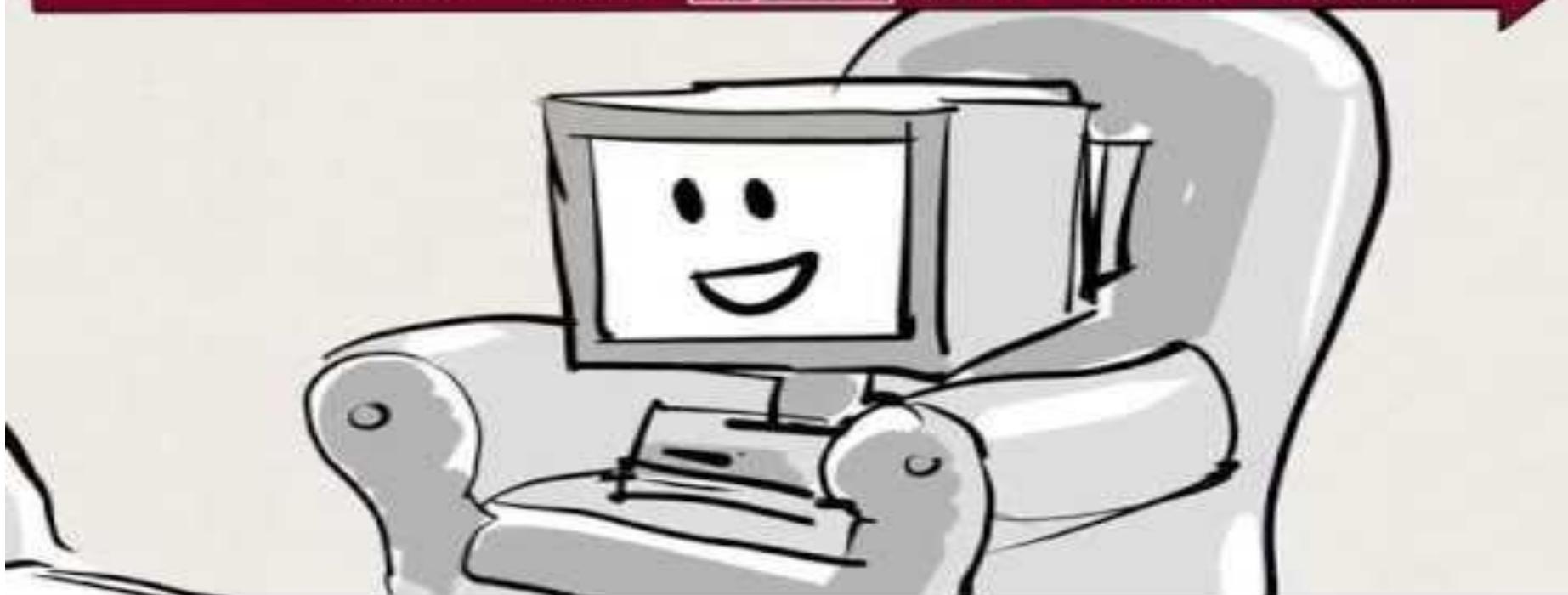
In WIRED's tests, Google Photos did identify some primates, but no gorillas like this one were to be found. RICK MADONIK/TORONTO STAR/GETTY IMAGES

<https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/>



# A brief history of AI

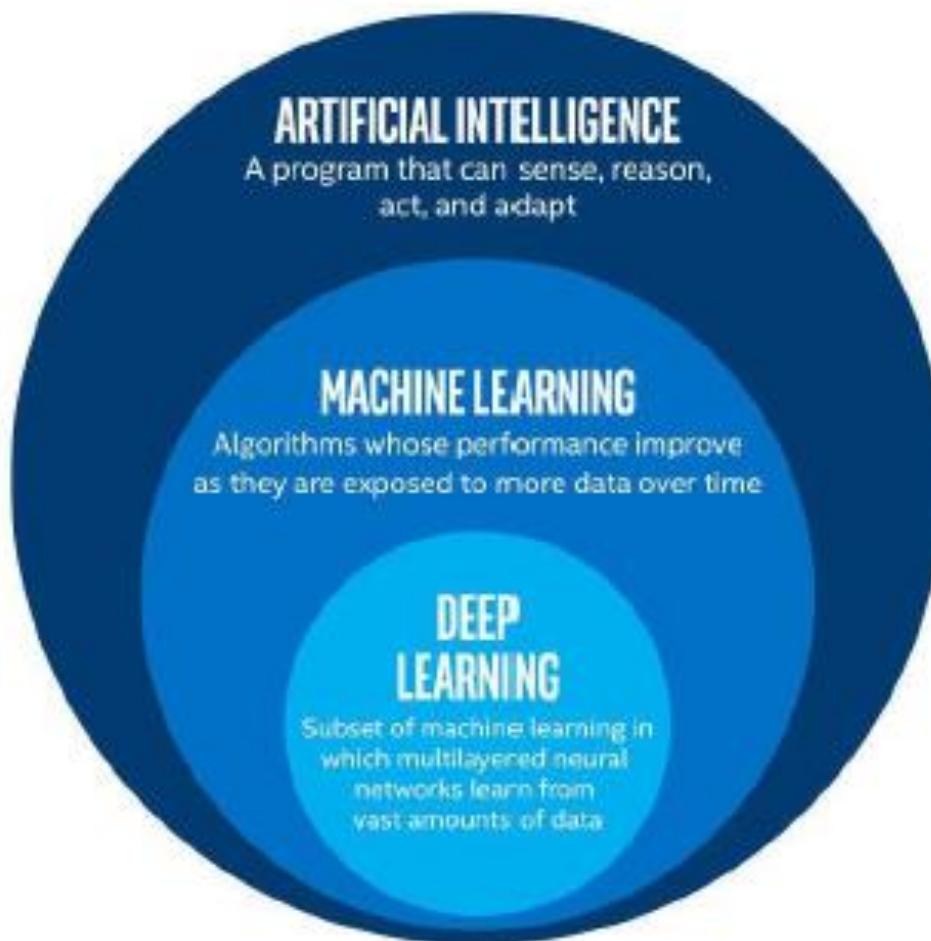
1950 1955 **1960** 1970 1980 1990





# Machine Learning

- These programs learn from repeatedly seeing data, rather than being explicitly programmed by humans



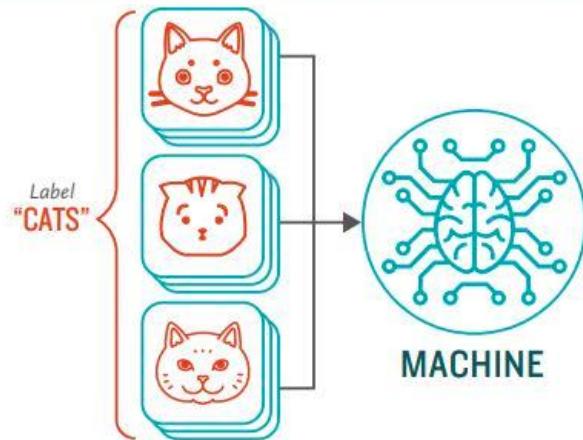


# Supervised Learning

## How **Supervised** Machine Learning Works

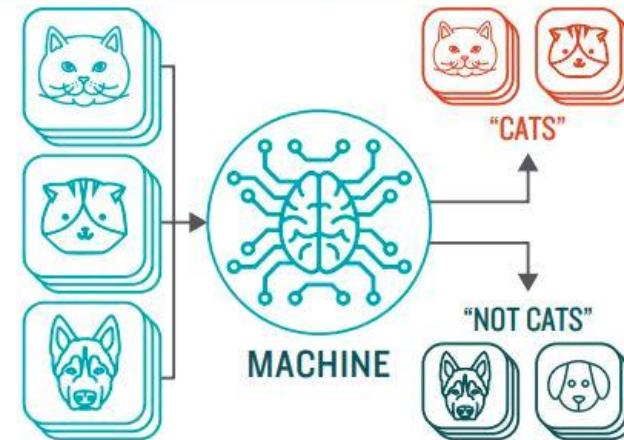
### STEP 1

Provide the machine learning algorithm categorized or "labeled" input and output data from to learn

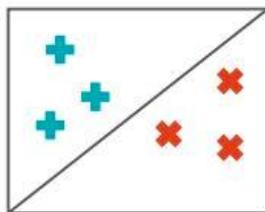


### STEP 2

Feed the machine new, unlabeled information to see if it tags new data appropriately. If not, continue refining the algorithm

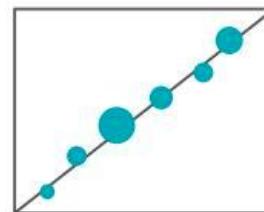


### TYPES OF PROBLEMS TO WHICH IT'S SUITED



#### CLASSIFICATION

Sorting items into categories



#### REGRESSION

Identifying real values (dollars, weight, etc.)

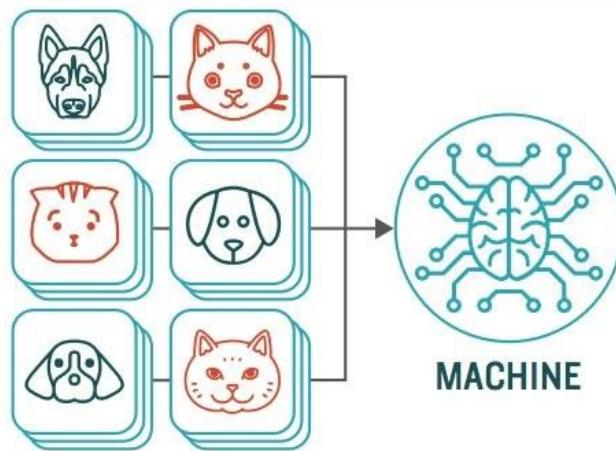


# Unsupervised Learning

## How **Unsupervised** Machine Learning Works

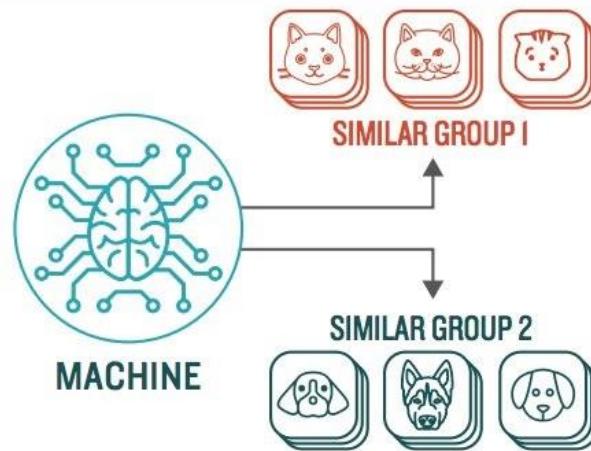
### STEP 1

Provide the machine learning algorithm uncategorized, unlabeled input data to see what patterns it finds

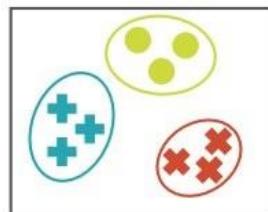


### STEP 2

Observe and learn from the patterns the machine identifies



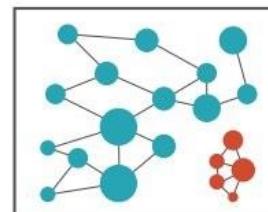
## TYPES OF PROBLEMS TO WHICH IT'S SUITED



### CLUSTERING

Identifying similarities in groups

*For Example:* Are there patterns in the data to indicate certain patients will respond better to this treatment than others?



### ANOMALY DETECTION

Identifying abnormalities in data

*For Example:* Is a hacker intruding in our network?



# Machine Learning

---

- **Two main types of learning**
  - Supervised Learning
    - Data points have known outcome
    - Goal is to make predictions - Classification and Regression
  - Unsupervised Learning
    - Data points have unknown outcome
    - Goal is to find structure within the data – Clustering
- **Other types of learning**
  - Reinforcement Learning
  - Genetic Algorithm



# Machine Learning

- Applications in our daily lives

Spam Filtering

Web Search

Postal Mail Routing

Fraud Detection

Movie  
Recommendations

Vehicle Driver  
Assistance

Web Advertisements

Social Networks

Speech Recognition



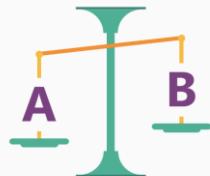
# 5 fundamental questions

**Is this weird?  
(Anomaly detection)**



Is this pressure gauge reading normal?  
Is this message from the internet typical?

**Is this A or B?  
(Classification)  
(discrete values)**



Will this tire fail in the next 1,000 miles: Yes or no?  
Which brings in more customers: a \$5 coupon or a 25% discount?

**How many?  
How Much?  
(Regression)  
(Continuous)**



What will the temperature be next Tuesday?  
What will my fourth quarter sales be?

**How is this organized?  
(Clustering)**



Which viewers like the same types of movies?  
Which printer models fail the same way?

**What should I do?  
(Reinforce Learning)**

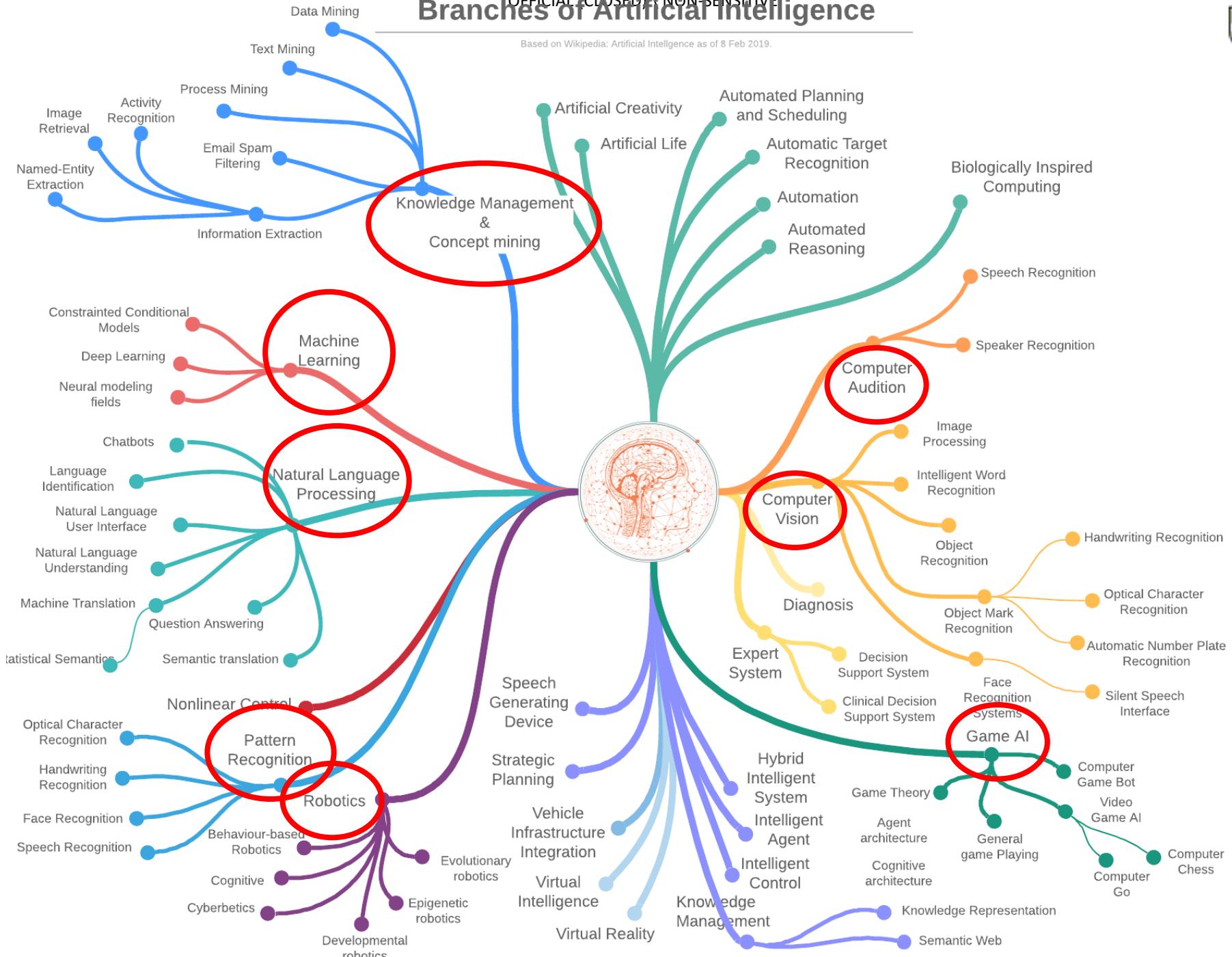


If I'm a self-driving car: At a yellow light, brake or accelerate?  
For a robot vacuum: Keep vacuuming, or go back to the charging station?



# OFFICIAL (CLOSED) X NON-SENSITIVE Branches of Artificial Intelligence

Based on Wikipedia: Artificial Intelligence as of 8 Feb 2019.





# Machine Learning Example

- Suppose you wanted to identify fraudulent credit card transactions.
- You could define features to be:
  - Transaction time
  - Transaction amount
  - Transaction location
  - Category of purchase
- The algorithm could learn what feature combinations suggest unusual activity.





Tuesday, April 06, 2021

# AI tapped to predict cheaters' next move

David Sun

Bank fraud experts are using artificial intelligence (AI) to outsmart scammers who think they can fool financial institutions into believing they are customers and gain access to funds in accounts.

They are harnessing the technology to monitor scammers' behaviour to predict their next move.

With the assistance of machine learning and data analytics, banks and the police have built systems to monitor transactions on unprecedented scale.

Singapore Police Force's e-Security Centre (ASC) said it has tested robotic process automation to its operations.

Creates a standardised format for sharing of information on various accounts via APIs, enabling them to take steps taken by bad actors, says Royston Soon, vice-president of risk management. And its systems are able to identify transaction scams.

"Using a machine-learning model, we are able to then rate the probability of the next transaction being fraudulent," he added.

"If the system detects a behaviour on a new device, an alert will be triggered, and we can use this to identify money mules. We are then able to take action on such accounts even before the funds get transferred out."

Soon said OCBC launched its surveillance system in 2016. The system monitors online transactions and gives each a risk score. High-risk ones trigger an alert, which is then reviewed by a fraud analyst.

## FRAUD DETECTION

Even if you are using spoof technology, this system is able to identify your true IP and location. Fraudsters typically use one or a few devices to access multiple money mule accounts, which will then be flagged to our fraud analysts.



**MR ROYSTON SOON**, vice-president of fraud risk management at OCBC Bank.

He said that a machine-learning

The system also captures behavioural biometrics, noting a user's typical finger movements and clicks on a device.

Mr Soon said the system would be able to capture one's mouse movements, typing speed and how a user navigates, backspaces and uses autofill components.

analysts."

The system also captures behavioural biometrics, noting a user's typical finger movements and clicks on a device.

Mr Soon said the system would be able to capture one's mouse movements, typing speed and how a user navigates, backspaces and uses autofill components.

"We are able to then differenti-

ate whether or not the one interacting on our platform is really our client," he noted.

Fraud analysts can then account takeover scenario to determine if it is a case of an account being used by a money mule.

He said the system is constantly learning as new data points are added.

Together with the information from the ASC, the bank is able to create what is essentially a scam-hunting sentinel or form.

DBS Bank also has its own detection system that uses machine learning and AI.

Mr Elvin Lim, head of group investigation at the bank, said: "We rely on machine learning and AI to basically identify transaction patterns of usage of customers' accounts, and using all the data points to tell us the story about whether this is a real customer or a potential scammer."

He pointed out that this network link analysis is like a spider web.

When one account gets hit, they analyse the kind of transactions that happen through that account.

The data is then processed in a pictorial form to link other frauds together and provide a comprehensive picture of the operations. Machine learning is also potentially fraudulently used, said UOB, said Mr Lim. In addition to investigating integrated Fraud department, these systems have been well received by the banking community as scammers change tactics, he added.

"This is why we hope to strengthen our collaboration with the other banks through the ASC," he said. "A combined approach with the whole banking community is needed so that we can all harness and build on technology together to complement our efforts to stop crime."

**Mr Elvin Lim**, head of group investigation at the bank, said: "We rely on machine learning and AI to basically identify transaction patterns of usage of customers' accounts, and using all the data points to tell us the story about whether this is a real customer or a potential scammer."

"This is why we hope to strengthen our collaboration with the other banks through the ASC," he said. "A combined approach with the whole banking community is needed so that we can all harness and build on technology together to complement our efforts to stop crime."

"Using a machine-learning model, we are able to then rate the probability of the next transaction being fraudulent," he added.

"If the system detects a behaviour on a new device, an alert will be triggered, and we can use this to identify money mules. We are then able to take action on such accounts even before the funds get transferred out."

The system monitors online transactions and gives each a risk score. High-risk ones trigger an alert, which is then reviewed by a fraud analyst.



# Machine Learning Limitations

---

- Suppose you wanted to determine if an image is of a cat or a dog.
- What features would you use?
- This is where **Deep Learning** can come in.



*Dog and cat recognition*

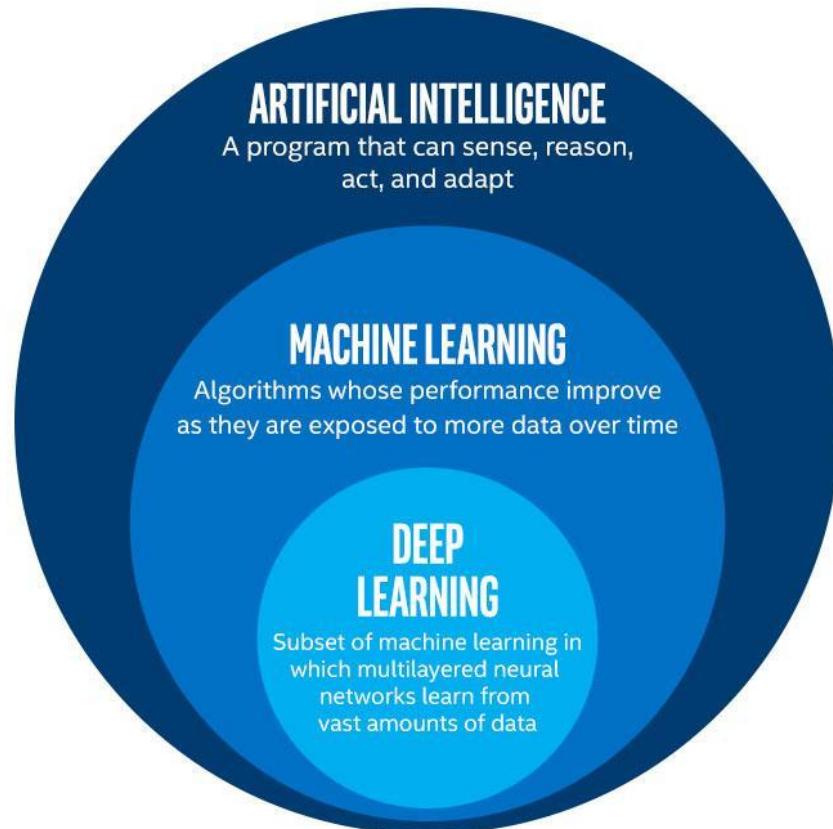


# What is deep learning?

## Deep Learning

“Machine learning that involves using very complicated models called “deep neural networks”.”  
(Intel)

*Models* determine best representation of original data; in classic machine learning, humans must do this.

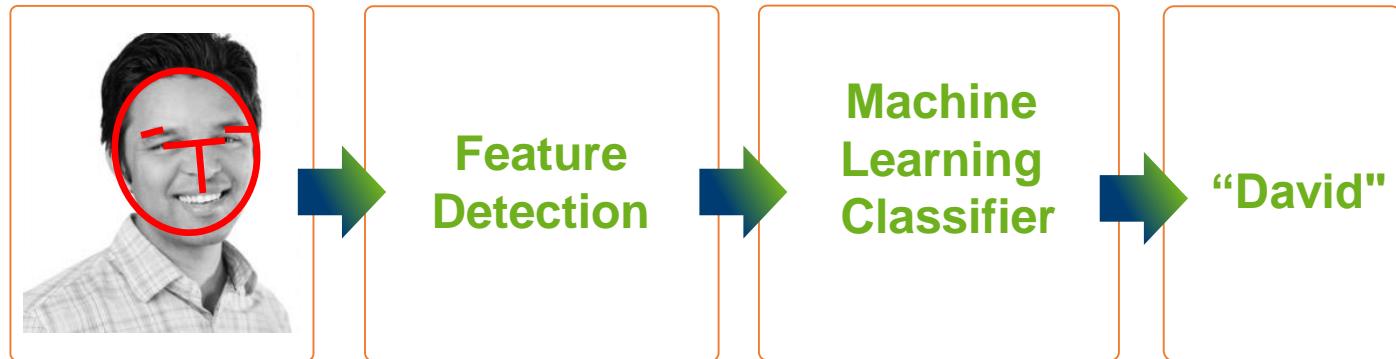




# Deep Learning Example

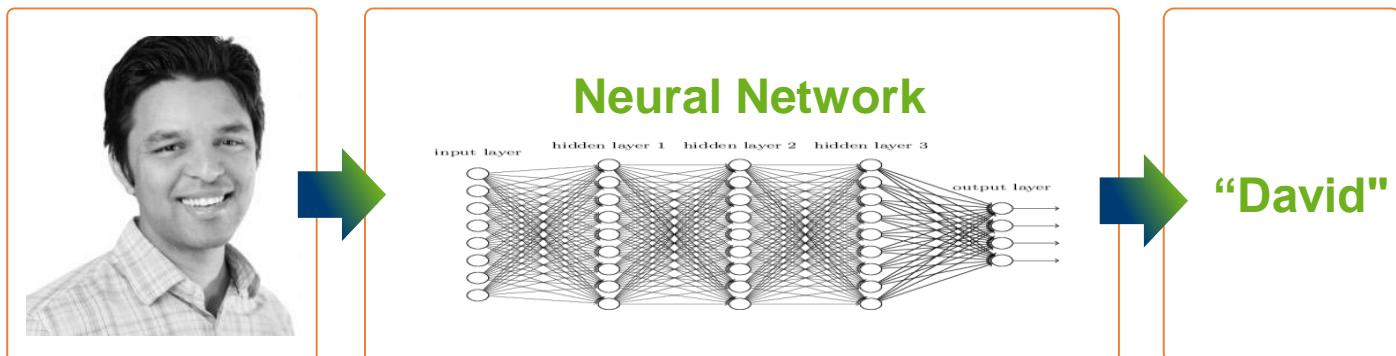
## Classic Machine Learning

Step 1: Determine features.  
Step 2: Feed them through model.



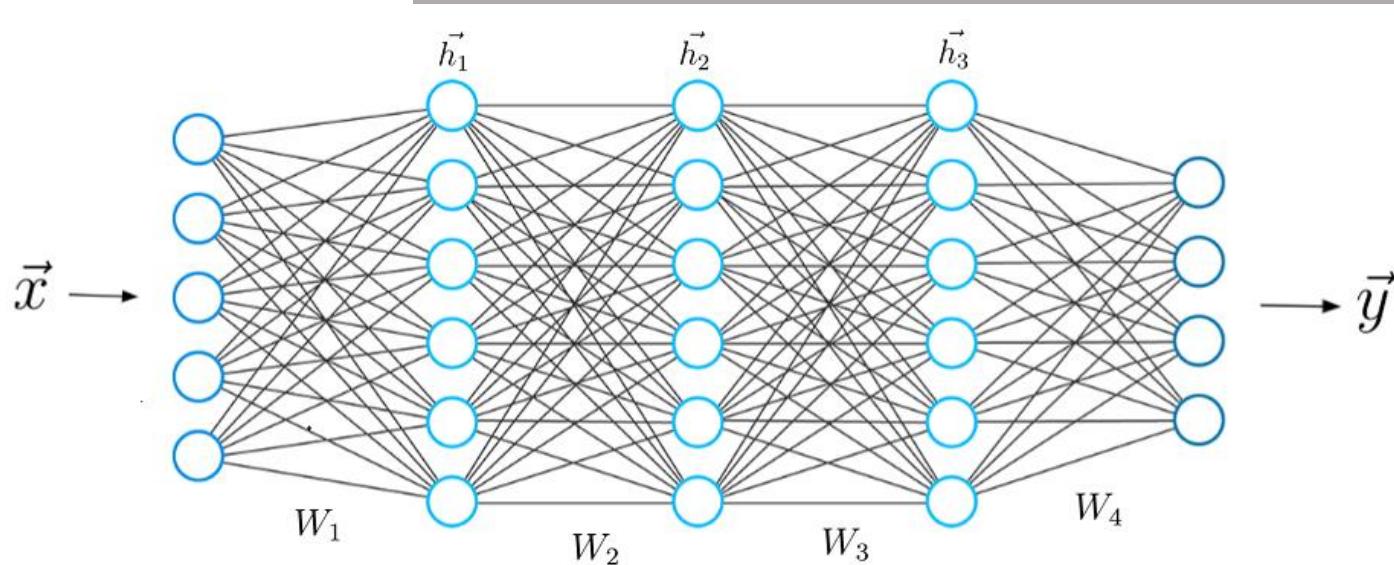
## Deep Learning

Steps 1 and 2 are combined into 1 step.

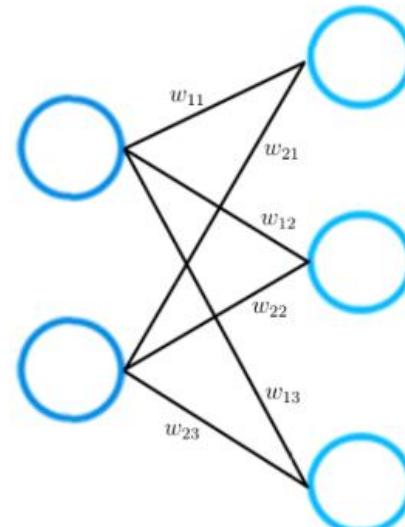




# Neural Networks

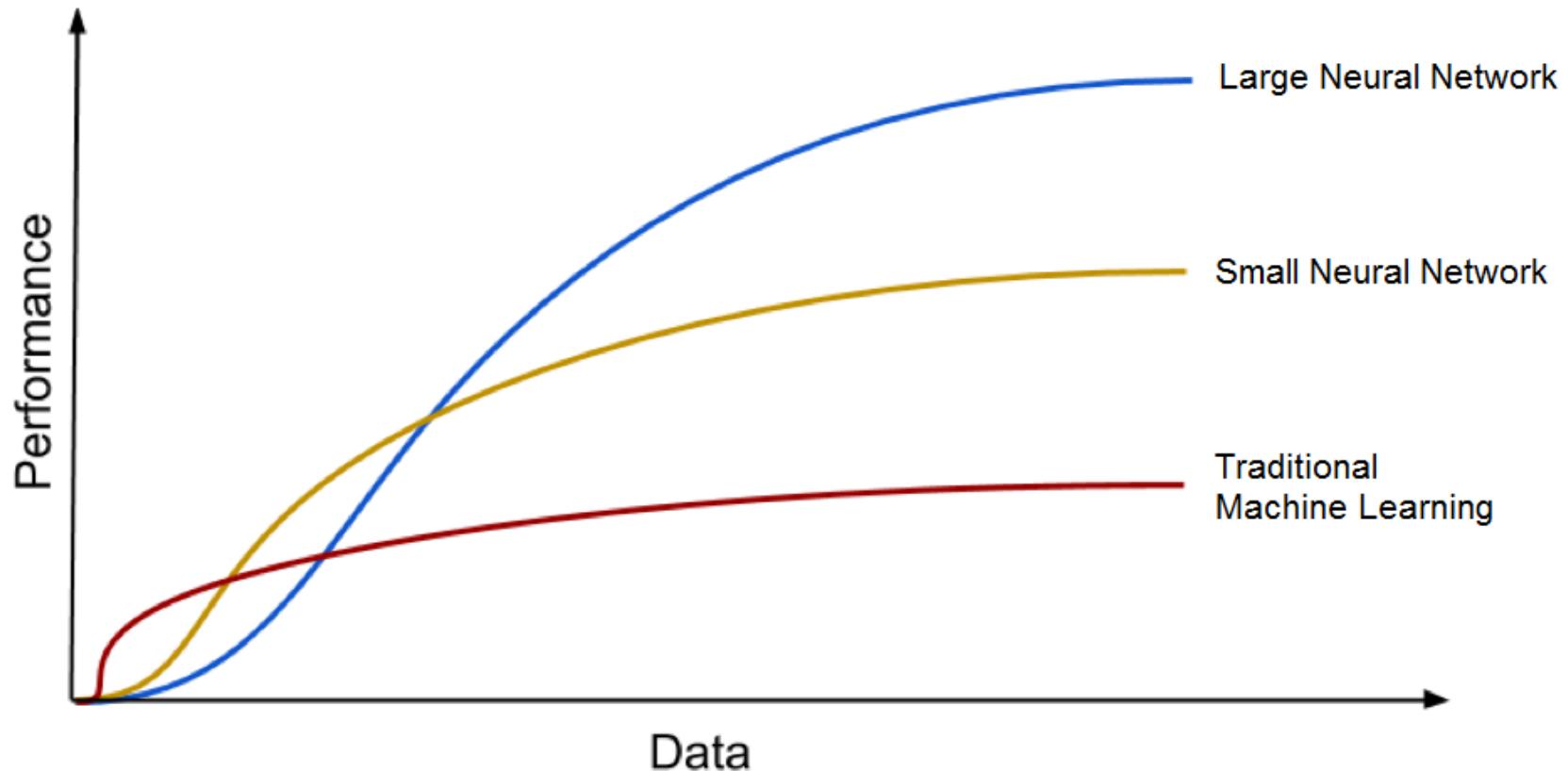


The challenge in training a neural networks is finding a set of weights the give the most accurate output.





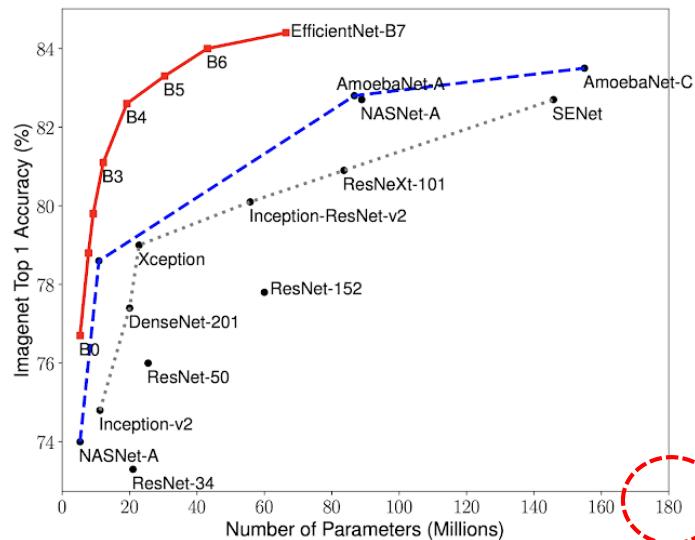
# Performance



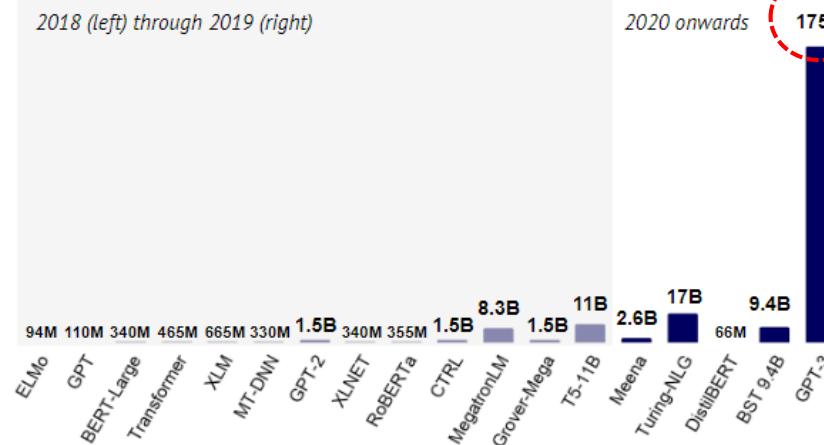
Deep Learning Algorithms get better with the increasing amount of data.



# Size



► Huge models, large companies and massive training costs dominate the hottest area of AI today, NLP.



Note: The number of parameters indicates how many different coefficients the algorithm optimizes during the training process.



# Deep Learning in Action

[bit.ly/google\\_teachable](https://bit.ly/google_teachable)

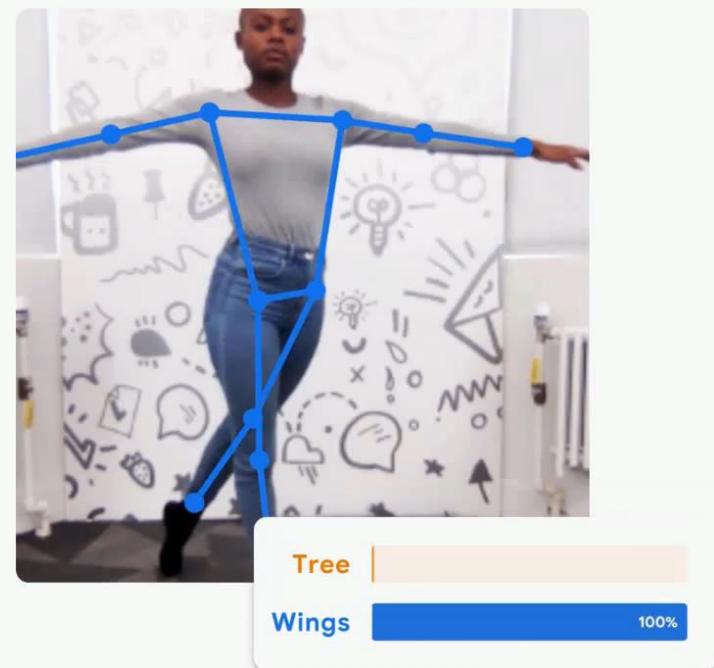
## Teachable Machine



Train a computer to recognize your own images, sounds, & poses.

A fast, easy way to create machine learning models for your sites, apps, and more – no expertise or coding required.

Get Started

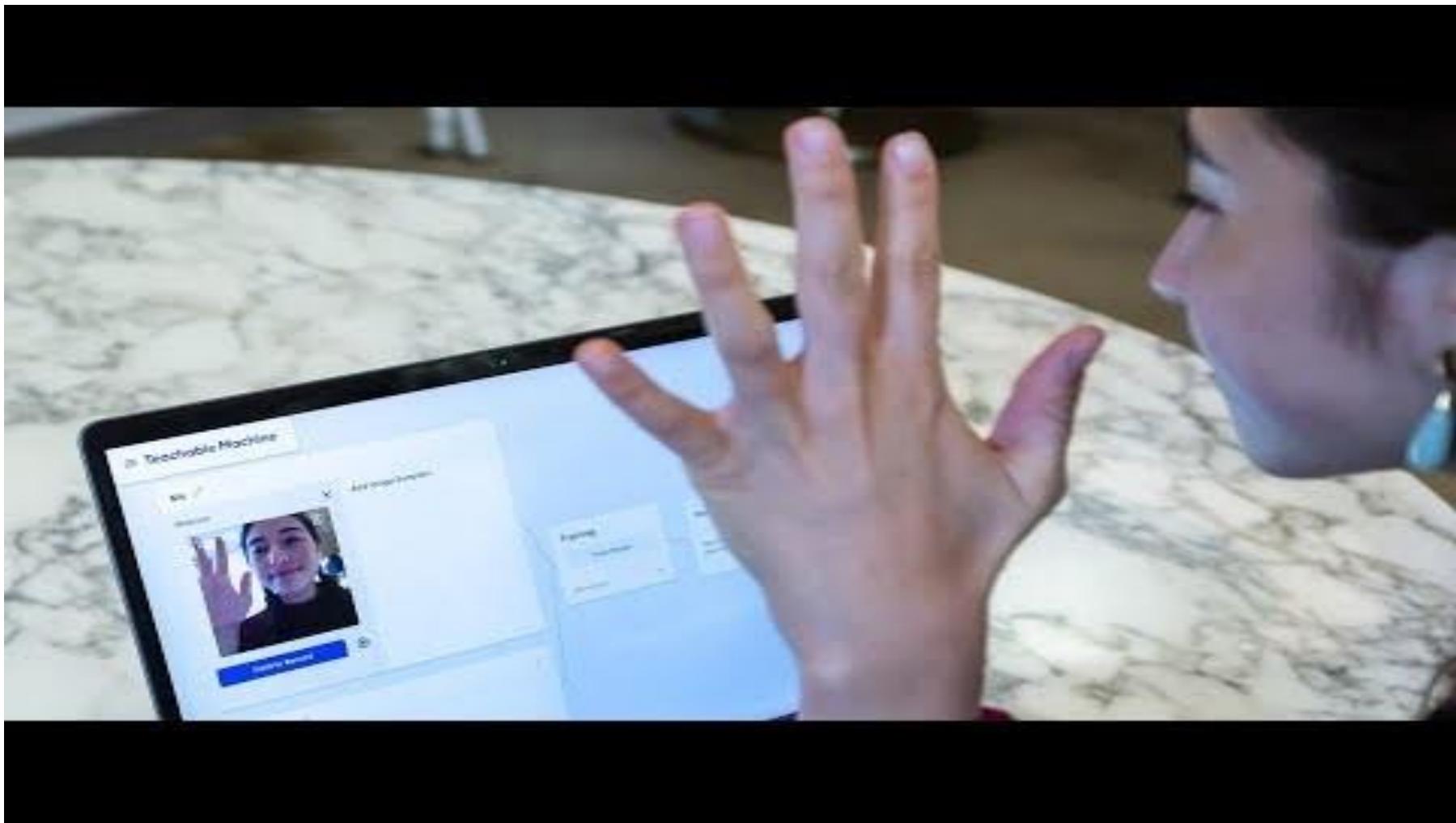


Optional Activity



# Teachable Machine

---



<https://teachablemachine.withgoogle.com/>



# 15 Mins Break

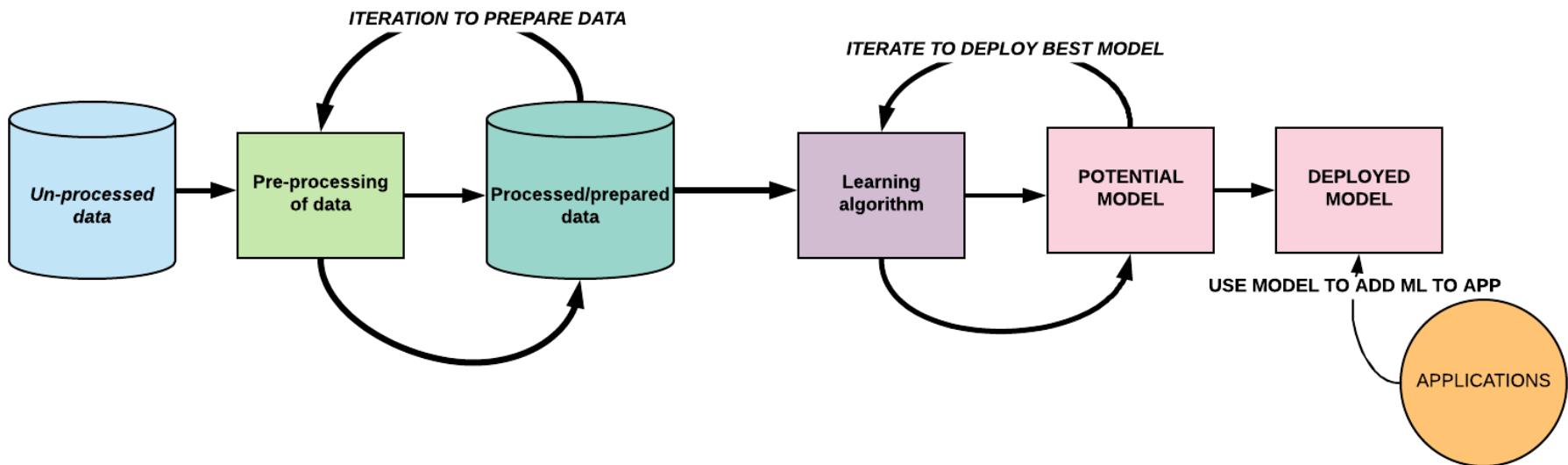


[bit.ly/top10\\_2020](https://bit.ly/top10_2020)





# Machine Learning workflow

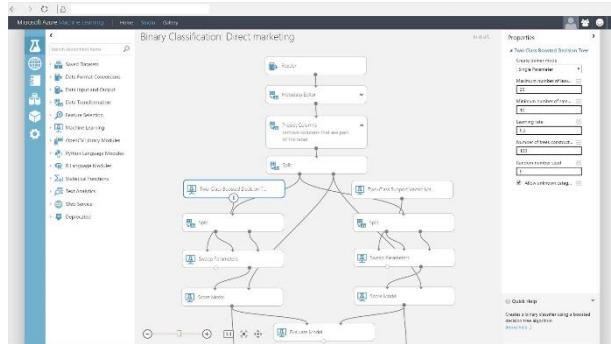


Ref: <https://cloudacademy.com/blog/what-is-azure-machine-learning/>

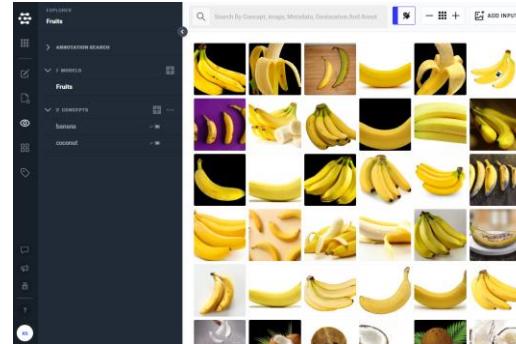


# Code-Free Machine Learning tools

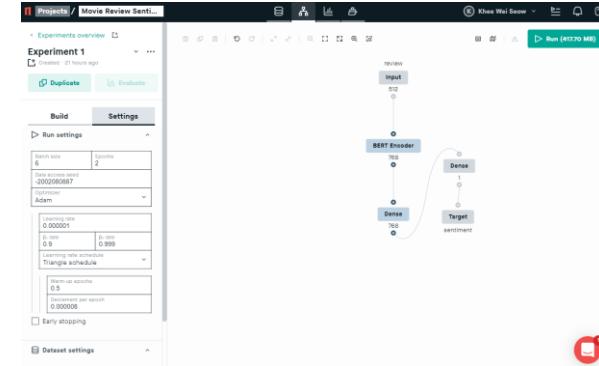
Microsoft Azure  
Machine Learning Studio  
(Classic)



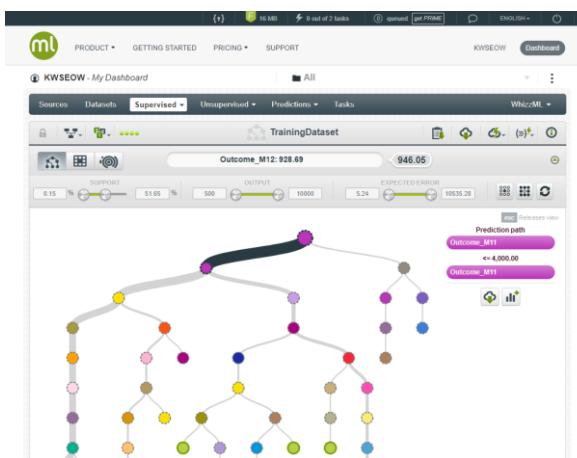
Clarifai



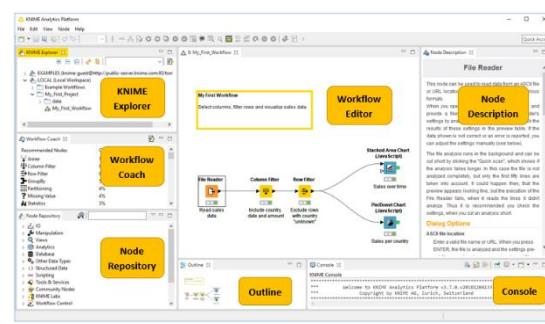
Peltarion



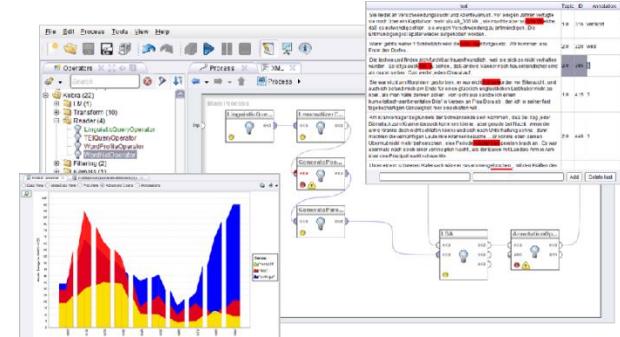
bigml



KNIME



Rapidminer





# Activity 1 – First Machine Learning with Azure

- Automobile Price Prediction

Given some features of a car, e.g. engine capacity, no of doors, horsepower, predict the selling price



**Step 1:**  
Watch and listen to the instructor's demonstration



30 mins

symboling	normalized_msrp	make	fuel-type	aspiration	num-of-drives	body-style	drive-wheel	engine-loc	wheel-base	length	width	height	curb-weight	engine-type	num-of-cyl	engine-size	fuel-system	bore	stroke
3 ?	10.8	alfa-romeo	gas	std	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi	3.47	2.68
3 ?	10.8	alfa-romeo	gas	std	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi	3.47	2.68
1 ?	10.8	alfa-romeo	gas	std	two	hatchback	rwd	front	94.5	171.2	65.5	52.4	2823	ohcv	six	152	mpfi	2.68	3.47
2	164	audi	gas	std	four	sedan	fwd	front	99.8	176.6	66.2	54.3	2373	ohc	four	109	mpfi	3.19	3.4
2	164	audi	gas	std	four	sedan	4wd	front	99.4	176.6	66.4	54.3	2824	dohc	five	136	mpfi	3.19	3.4
2 ?	164	audi	gas	std	two	sedan	fwd	front	99.8	177.3	66.3	53.1	2507	ohc	five	136	mpfi	3.19	3.4
1	158	audi	gas	std	four	sedan	fwd	front	105.8	192.7	71.4	55.7	2844	ohc	five	136	mpfi	3.19	3.4
1 ?	158	audi	gas	std	four	wagon	fwd	front	105.8	192.7	71.4	55.7	2954	ohc	five	136	mpfi	3.19	3.4
1	158	audi	gas	turbo	four	sedan	fwd	front	105.8	192.7	71.4	55.9	3086	ohc	five	131	mpfi	3.13	3.4
0 ?	192	audi	gas	turbo	two	hatchback	4wd	front	99.5	178.2	67.9	52	3053	ohc	five	131	mpfi	3.13	3.4
2	192	bmw	gas	std	two	sedan	rwd	front	101.2	176.8	64.8	54.3	2395	ohc	four	108	mpfi	3.5	2.8
0	192	bmw	gas	std	four	sedan	rwd	front	101.2	176.8	64.8	54.3	2395	ohc	four	108	mpfi	3.5	2.8
0	192	bmw	gas	std	two	sedan	rwd	front	101.2	176.8	64.8	54.3	2711	ohc	six	164	mpfi	3.2	3.3
0	192	bmw	gas	std	four	sedan	rwd	front	101.2	176.8	64.8	54.3	2765	ohc	six	164	mpfi	3.21	3.3
1 ?	192	bmw	gas	std	four	sedan	rwd	front	103.5	189	66.9	55.7	3055	ohc	six	164	mpfi	3.21	3.3
0 ?	209	bmw	gas	std	two	sedan	rwd	front	103.5	189	66.9	55.7	3230	ohc	six	209	mpfi	3.62	3.39
0 ?	209	bmw	gas	std	four	sedan	rwd	front	103.5	193.8	67.9	53.7	3380	ohc	six	209	mpfi	3.62	3.39
2	211	chevrolet	gas	std	two	hatchback	rwd	front	88.4	141.1	60.3	53.2	1488	i	three	61	2bbl	2.91	3.03
1	98	chevrolet	gas	std	two	hatchback	rwd	front	94.5	155.9	63.6	52	1874	ohc	four	90	2bbl	3.03	3.11
0	81	chevrolet	gas	std	four	sedan	rwd	front	94.5	158.8	63.6	52	196	ohc	four	90	2bbl	3.03	3.11
1	118	dodge	gas	std	two	hatchback	rwd	front	93.7	157.3	63.8	50.8	1876	ohc	four	90	2bbl	2.97	3.23
1	118	dodge	gas	std	two	hatchback	rwd	front	93.7	157.3	63.8	50.8	1876	ohc	four	90	2bbl	2.97	3.23
1	118	dodge	gas	turbo	two	hatchback	rwd	front	93.7	157.3	63.8	50.8	2116	ohc	four	98	mpfi	3.03	3.39
1	148	dodge	gas	std	four	hatchback	rwd	front	93.7	157.3	63.8	50.6	1967	ohc	four	90	2bbl	2.97	3.23
1	148	dodge	gas	std	four	sedan	rwd	front	93.7	157.3	63.8	50.8	1989	ohc	four	90	2bbl	2.97	3.23
1	148	dodge	gas	std	four	sedan	rwd	front	93.7	157.3	63.8	50.6	1989	ohc	four	90	2bbl	2.97	3.23
1	148	dodge	gas	turbo	?	sedan	rwd	front	93.7	157.3	63.8	50.6	2191	ohc	four	98	mpfi	3.03	3.39
1	110	dodge	gas	std	four	wagon	rwd	front	103.3	174.6	64.6	59.8	2355	ohc	four	122	2bbl	3.34	3.46
3	145	dodge	gas	turbo	two	hatchback	rwd	front	95.9	172.2	66.3	50.2	2811	ohc	four	156	mpfi	3.6	3.9
2	137	honda	gas	std	two	hatchback	rwd	front	86.6	144.6	63.9	50.8	1713	ohc	four	92	1bbl	2.91	3.41
2	137	honda	gas	std	two	hatchback	rwd	front	86.6	144.6	63.9	50.8	1819	ohc	four	92	1bbl	2.91	3.41
1	101	honda	gas	std	two	hatchback	rwd	front	93.7	150	64	52.6	1837	ohc	four	79	1bbl	2.91	3.07
1	101	honda	gas	std	two	hatchback	rwd	front	93.7	150	64	52.6	1948	ohc	four	92	1bbl	2.91	3.41
0	110	honda	gas	std	four	sedan	rwd	front	96.5	163.4	64	54.5	2010	ohc	four	92	1bbl	2.91	3.41
0	78	honda	gas	std	four	wagon	rwd	front	96.5	151.1	63.9	58.3	2024	ohc	four	92	1bbl	2.92	3.41
0	106	honda	gas	std	two	hatchback	rwd	front	96.5	167.5	65.2	53.3	2236	ohc	four	110	1bbl	3.15	3.58
0	85	honda	gas	std	four	sedan	rwd	front	96.5	175.4	65.2	54.1	2306	ohc	four	110	1bbl	3.15	3.58
0	85	honda	gas	std	four	sedan	rwd	front	96.5	175.4	65.2	54.1	2372	ohc	four	110	1bbl	3.15	3.58
0	85	honda	gas	std	four	sedan	rwd	front	96.5	175.4	65.2	54.1	2406	ohc	four	110	1bbl	3.15	3.58
1	207	honda	gas	std	two	sedan	rwd	front	96.5	169.1	65.6	51	2293	ohc	four	110	2bbl	3.15	3.58
0 ?	110	isuzu	gas	std	four	sedan	rwd	front	94.3	170.7	61.8	53.5	2337	ohc	four	111	2bbl	3.31	3.23
1 ?	110	isuzu	gas	std	two	sedan	rwd	front	94.3	170.7	61.8	53.5	2347	ohc	four	90	2bbl	3.03	3.11
0 ?	110	isuzu	gas	std	four	sedan	rwd	front	94.5	155.9	63.6	52	1909	ohc	four	90	2bbl	3.03	3.11
2 ?	110	isuzu	gas	std	two	hatchback	rwd	front	96	172.6	65.2	51.4	2734	ohc	four	119	spfi	3.43	3.23
0	145	jaguar	gas	std	four	sedan	rwd	front	113	199.6	69.6	52.8	406	dohc	six	258	mpfi	3.63	4.17
0 ?	145	jaguar	gas	std	four	sedan	rwd	front	113	199.6	69.6	52.8	406	dohc	six	258	mpfi	3.63	4.17
0 ?	104	mazda	gas	std	two	hatchback	rwd	front	93.1	159.1	64.2	54.1	1890	ohc	four	91	2bbl	3.03	3.15
1	104	mazda	gas	std	two	hatchback	rwd	front	93.1	159.1	64.2	54.1	1900	ohc	four	91	2bbl	3.03	3.15
1	113	mazda	gas	std	four	sedan	rwd	front	93.1	166.8	64.2	54.1	1945	ohc	four	91	2bbl	3.03	3.15
1	113	mazda	gas	std	two	hatchback	rwd	front	93.1	166.8	64.2	54.1	1950	ohc	four	91	2bbl	3.03	3.15
3	150	mazda	gas	std	two	hatchback	rwd	front	95.3	169	65.7	49.6	2380	rotor	two	70	4bbl	?	?
3	150	mazda	gas	std	two	hatchback	rwd	front	95.3	169	65.7	49.6	2385	rotor	two	70	4bbl	?	?
3	150	mazda	gas	std	two	hatchback	rwd	front	95.3	169	65.7	49.6	2500	rotor	two	80	mpfi	?	?
1	129	mazda	gas	std	two	hatchback	rwd	front	98.8	177.8	66.5	53.7	2385	ohc	four	122	2bbl	3.39	3.39
0	115	mazda	gas	std	four	sedan	rwd	front	98.8	177.8	66.5	53.7	2386	ohc	four	122	2bbl	3.39	3.39
0	115	mazda	diesel	std	?	sedan	rwd	front	98.8	177.8	66.5	55.5	2443	ohc	four	122	idi	3.39	3.39

**Individual Activity**

45 mins





# 60 mins Lunch Break

**Some interesting videos**

<https://www.youtube.com/watch?v=bmNaLtC6vkU>

[https://www.youtube.com/watch?v=Nnf8P5A\\_saE](https://www.youtube.com/watch?v=Nnf8P5A_saE)

LUNCH BREAK





# Recap

## Machine Learning in ML Studio

### Anomaly Detection

- One-class Support Vector Machine
- Principal Component Analysis-based Anomaly Detection
- Time Series Anomaly Detection\*

### Classification

#### Two-class Classification

- Averaged Perceptron
- Bayes Point Machine
- Boosted Decision Tree
- Decision Forest
- Decision Jungle
- Logistic Regression
- Neural Network
- Support Vector Machine

#### Multi-class Classification

- Decision Forest
- Decision Jungle
- Logistic Regression
- Neural Network
- One-vs-all

### Clustering

- K-means Clustering

### Recommendation

- Matchbox Recommender

### Regression

- Bayesian Linear Regression
- Boosted Decision Tree
- Decision Forest
- Fast Forest Quantile Regression
- Linear Regression
- Neural Network Regression
- Ordinal Regression
- Poisson Regression

### Statistical Functions

- Descriptive Statistics
- Hypothesis Testing T-Test
- Linear Correlation
- Probability Function Evaluation

### Text Analytics

- Feature Hashing
- Named Entity Recognition
- Vowpal Wabbit

### Computer Vision

- OpenCV Library

<https://studio.azureml.net>

Guest Access Workspace: Free trial access without logging in.

Free Workspace: Free persisted access, no Azure subscription needed.

Standard Workspace: Full access with SLA under an Azure subscription.

Cross browser drag & drop ML workflow designer.  
Zero installation needed.

Import Data

Preprocess

### Unlimited Extensibility

- R Script Module
- Python Script Module
- Custom Module
- Jupyter Notebook

Built-in ML Algorithms

Split Data

Train Model

Score Model

### Training Experiment

### One-click Operationalization

Predictive Experiment

### Make Prediction with Elastic APIs

- Request-Response Service (RRS)
- Batch Execution Service (BES)
- Retraining API

### Data Source

- Azure Blob Storage
- Azure SQL DB
- Azure SQL DW\*
- Azure Table
- Desktop Direct Upload
- Hadoop Hive Query
- Manual Data Entry
- OData Feed
- On-prem SQL Server\*
- Web URL (HTTP)

### Data Format

- ARFF
- CSV
- SVMLight
- TSV
- Excel
- ZIP

### Data Preparation

- Clean Missing Data
- Clip Outliers
- Edit Metadata
- Feature Selection
- Filter
- Learning with Counts
- Normalize Data
- Partition and Sample
- Principal Component Analysis
- Quantize Data
- SQLite Transformation
- Synthetic Minority Oversampling Technique

### Enterprise Grade Cloud Service

- SLA: 99.95% Guaranteed Up-time
- Azure AD Authentication
- Compute at Large Scale
- Multi-geo Availability
- Regulatory Compliance\*

### Community

- Gallery (<http://gallery.azureml.net>)
- Samples & Templates
- Workspace Sharing and Collaboration
- Live Chat & MSDN Forum Support

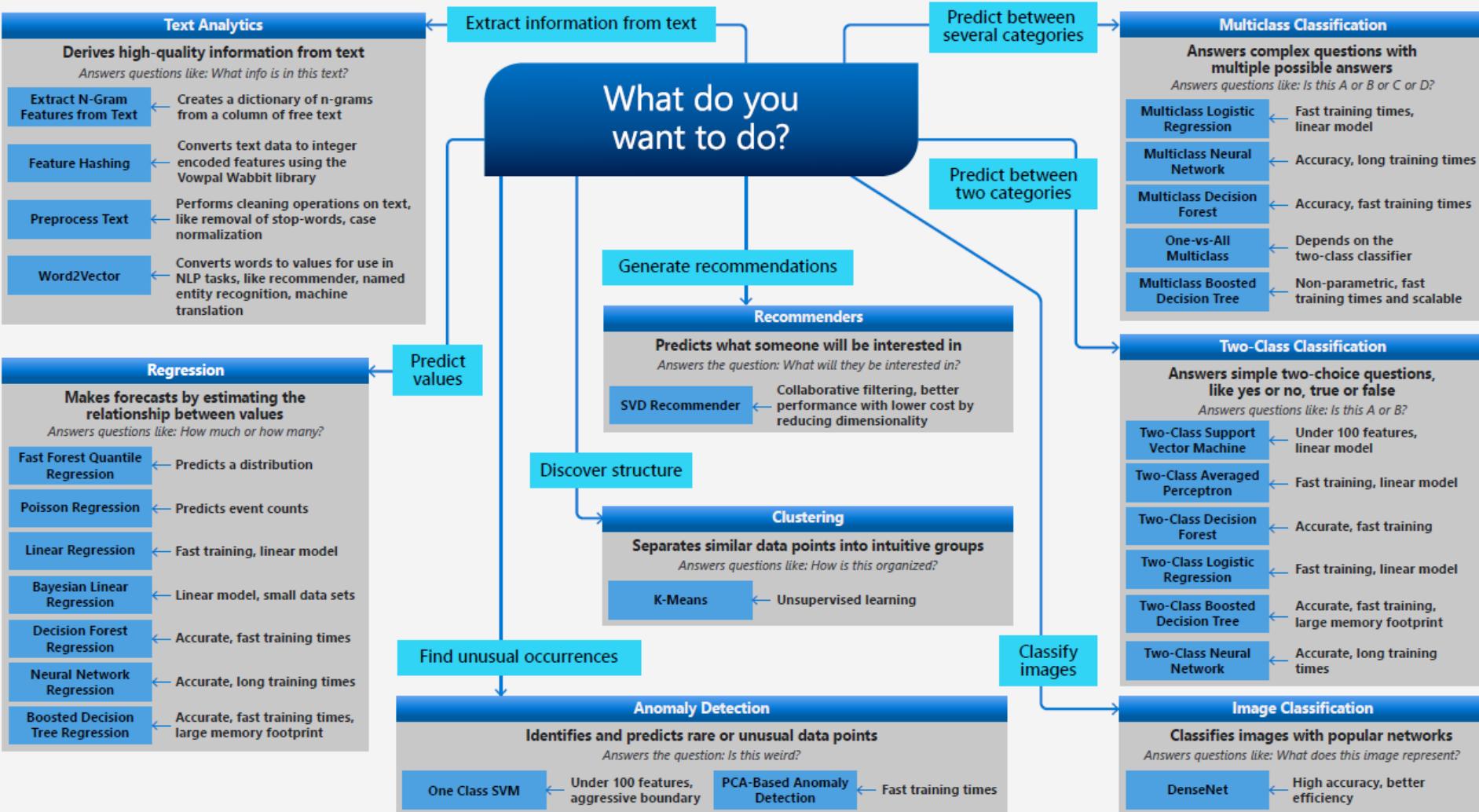
\* Feature Coming Soon



# Azure ML Algorithm Cheat Sheet

## Microsoft Azure Machine Learning Algorithm Cheat Sheet

This cheat sheet helps you choose the best machine learning algorithm for your predictive analytics solution. Your decision is driven by both the nature of your data and the goal you want to achieve with your data.





# Activity 2

- Deploying your experiment as a Web Service & Make Prediction using Excel

	A	B	C	D	E	F	G	H	I	J	
1	symboling	normalized	make	fuel-type	aspiration	num-of-doors	body-style	drive-wheel	engine-location	wheel-base	length
2	3	1	alfa-romero	gas	std	two	convertible	rwd	front	88.6	
3	3	1	alfa-romero	gas	std	two	convertible	rwd	front	88.6	
4	1	1	alfa-romero	gas	std	two	hatchback	rwd	front	94.5	
5	2	164	audi	gas	std	four	sedan	fwd	front	99.8	
6	2	164	audi	gas	std	four	sedan	4wd	front	99.4	
7											
8											
9											
10	make	body-style	wheel-base	engine-size	horsepower	peak-rpm	highway-mpg	price	Scored Labels		
11	alfa-romero	convertible	88.6	130	111	5000	27	13495	13498.476		
12	alfa-romero	convertible	88.6	130	111	5000	27	16500	13498.476		
13	alfa-romero	hatchback	94.5	152	154	5000	26	16500	14329.816		
14	audi	sedan	99.8	109	102	5500	30	13950	15696.502		
15	audi	sedan	99.4	136	115	5500	22	17450	17161.153		
16											

Azure Machine Learning interface showing the prediction process:

- 1. VIEW SCHEMA: My First Experiment [Predictive Exp.]
- 2. PREDICT: Input: input1 (Sheet1!A1:Z6, checked for headers), Output: output1 (Sheet1!A10, checked for headers). Predict button is highlighted.
- 3. ERRORS: None listed.

**Step 1:**  
Watch and listen to the  
instructor's demonstration



**Step 2:**  
Work through the activities

Individual Activity



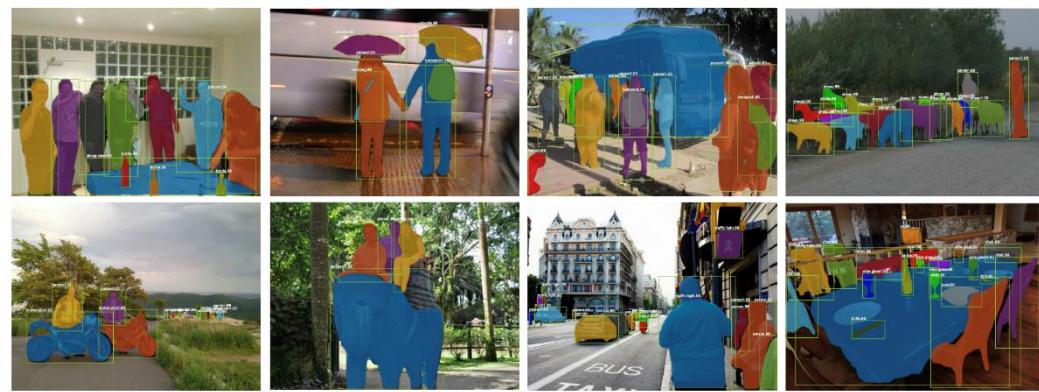
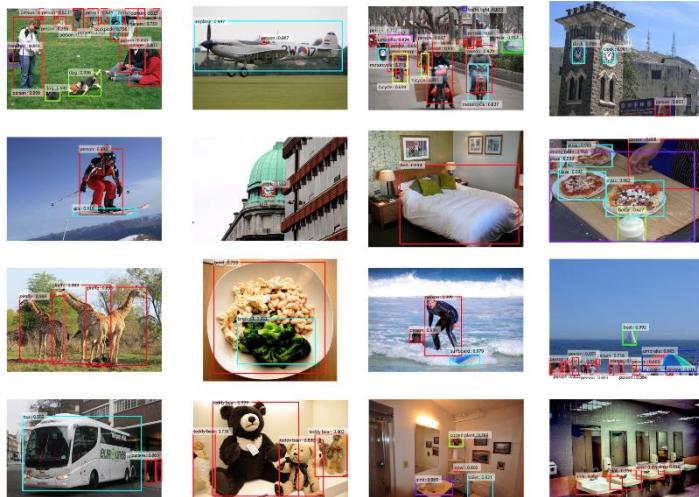
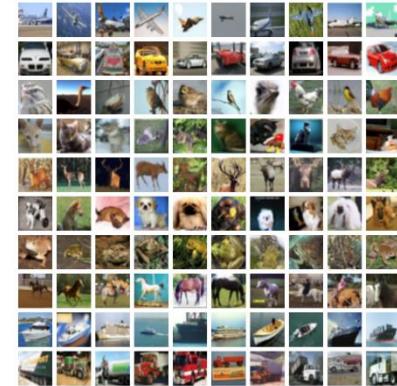


# Applications of Computer Vision

- Image Classification
- Image Classification With Localization
- Object Detection
- Object Segmentation



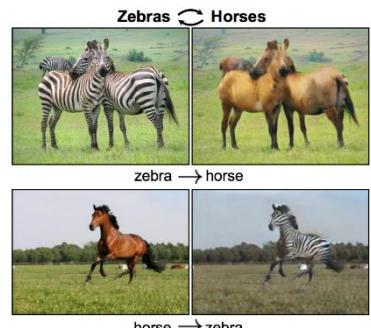
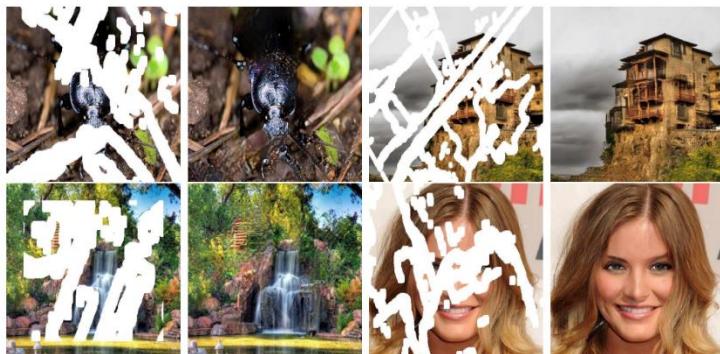
airplane  
automobile  
bird  
cat  
deer  
dog  
frog  
horse  
ship  
truck





# Applications of Computer Vision

- Image Style Transfer
- Image Colorization
- Image Reconstruction
- Image Super-Resolution
- Image Synthesis
- Other Problems





# Transfer Learning



Humans have an inherent ability to transfer knowledge across tasks.

What we acquire as knowledge while learning about one task, we utilize in the same way to solve related tasks.

The more related the tasks, the easier it is for us to transfer, or cross-utilize our knowledge.

Some simple examples would be,

- \* Know how to ride a motorbike → Learn how to ride a car
- \* Know how to play classic piano → Learn how to play jazz piano

- Models are difficult to train from scratch

- Huge datasets (like ImageNet)
- Long number of training iterations
- Very heavy computing machinery
- Time experimenting to get hyper-parameters just right

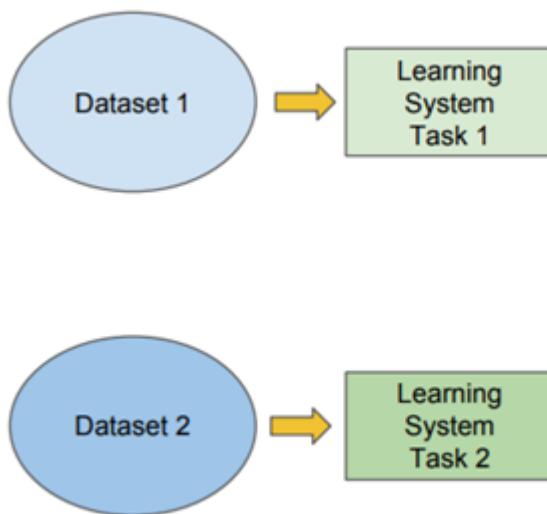


# Transfer Learning

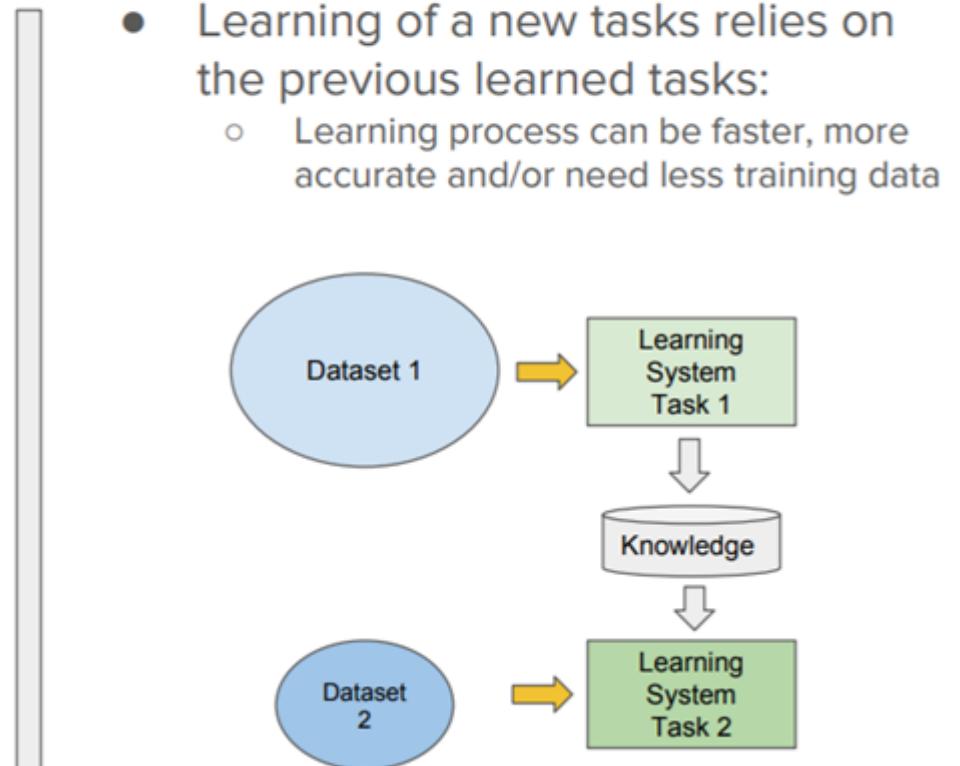
## Traditional ML

## vs Transfer Learning

- Isolated, single task learning:
  - Knowledge is not retained or accumulated. Learning is performed w.o. considering past learned knowledge in other tasks

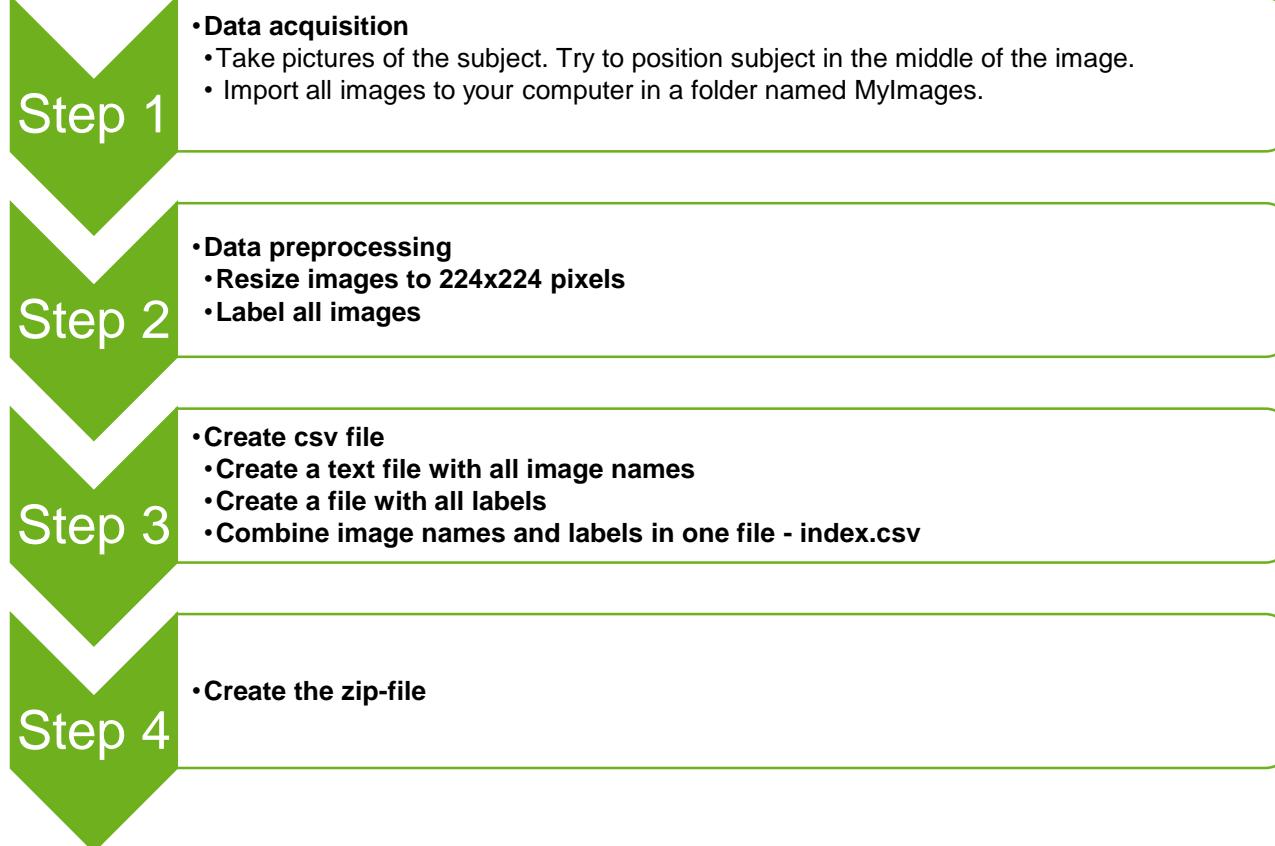


- Learning of a new tasks relies on the previous learned tasks:
  - Learning process can be faster, more accurate and/or need less training data





# Creating a new dataset





# Example

The diagram illustrates the structure of a car damage dataset. It shows a file tree, a preview of images, and a CSV metadata file.

**File Tree:**

- <> Dataset > Car damage dataset >
- Name
- image (highlighted with a red dashed box)
- test\_images
- index.csv (highlighted with a red dashed box)
- metadata.json

A green arrow points from the 'image' folder in the file tree to the 'image' column in the CSV table.

**Image Preview:**

<> Dataset > Car damage dataset > image

Search image

0.jpeg 1.jpeg 2.jpeg 3.jpeg 4.jpeg 5.jpeg

6.jpeg 7.jpeg 8.jpeg 9.jpeg 10.jpeg 11.jpeg

12.jpeg 13.jpeg 14.jpeg 15.jpeg 16.jpeg 17.jpeg

**CSV Metadata:**

A1	A	B	C
1	image	class	subset
2	image/0.jpeg	unknown	T
3	image/1.jpeg	head_lamp	T
4	image/2.jpeg	door_scratch	T
5	image/3.jpeg	head_lamp	T
6	image/4.jpeg	unknown	T
7	image/5.jpeg	unknown	T
8	image/6.jpeg	glass_shatter	T



# Activity 3 – Car Damage Classifier

A close-up photograph of a car's headlight that has been shattered into many pieces.	A photograph of a car's rear right light assembly, which appears to be broken or severely damaged.	A photograph of a car's side window that has been shattered, showing a large hole and broken glass shards.	A photograph of a car's white door panel with a visible horizontal scratch.
A photograph of a car's white door panel with a prominent dent.	A photograph of a dark-colored car's front bumper that has been dented.	A photograph of a white car's side bumper with a visible scratch.	A photograph of a silver SUV parked in a showroom setting.

**Step 1:**  
Watch and listen to the  
instructor's demonstration



20 mins

**Step 2:**  
- Do on your own



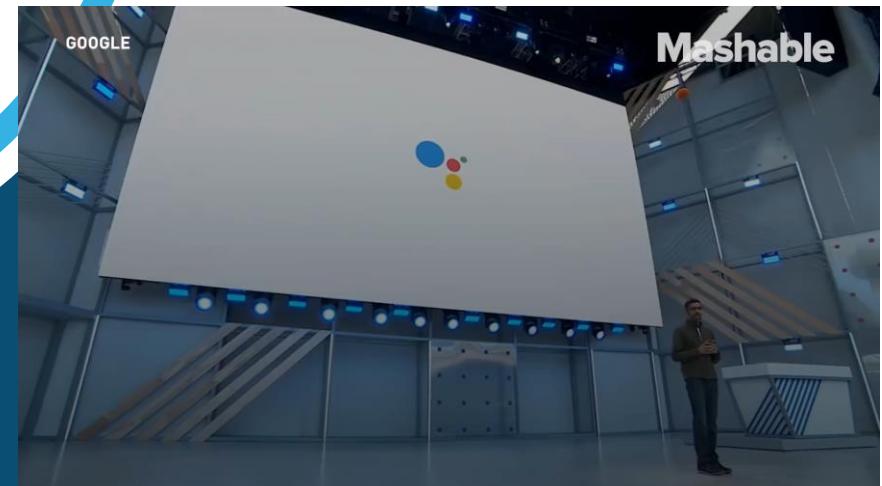
40 mins

**Individual Activity**



# 15 Mins Break

[bit.ly/google\\_duplex2019](https://bit.ly/google_duplex2019)





# Natural Language Processing

- Search Autocorrect and Autocomplete
- Language Translator
- **Social Media Monitoring**
- Chatbots
- **Survey Analysis**
- Targeted Advertising
- Hiring and Recruitment
- Voice Assistants
- Grammar Checkers
- Email Filtering





# Dataset

review	sentiment
Encoding Text	Encoding Binary Positive class positive
1 Hubert Selby Jr. gave us the book "Requiem For A Dream" and co-wrote the screenplay to Aronofsky's movie of it. That movie succeeded on every level by delivering an intimate, and unbiased portrait of the horrors of the characters lives and the vices that destroyed them. "Last Exit To Brooklyn" still has the vice and the multiple characters living sad lives, but it hardly does them the same justice Aronofsky did.  The film seems laughably anti-gay at times. Especially when in the film homosexuality equals death. One gay character gets stoned, is launched skyward by a speeding car, and lands dead on the pavement. Another is crucified and still more are simply beat up. Another exaggerated piece of shock value, that might actually hav...	negative
2 There are very few performers today who can keep me captivated throughout an entire film just by their presence. One of those few is Judy Davis, who has built a successful career out of creating characters that are headstrong in attitude but very vulnerable at heart. She takes roles that most other performers would treat melodramatically and adds a fiery, deeply emotional intensity that pulls attention away from everything else on the screen.  Her skills are well displayed in "High Tide," a film that matches her up a second time with director Gillian Armstrong, who gave Davis her first major success with "My Brilliant Career." In that film, Davis played a young woman who was determined to make it in the world, despite the suffocat...	positive
3 C'mon guys some previous reviewers have nearly written a novel commenting on this episode. It's just an old 60's TV show ! This episode of Star Trek is notable	negative

## Information

Creator

Review, Sentiment

Rows

25 000

Size

13 MB

Categories

Text, Classification



# Activity 4 - Creating a Sentiment Analyser



## About this dataset

This dataset contains textual movie reviews from IMDB users, together with the rating (simplified as positive or negative) that the user gave to the movie.

## Inspiration

Use this dataset to predict a simple positive or negative category from paragraph-sized text data.

### Step 1:

Watch and listen to the instructor's demonstration



20 mins

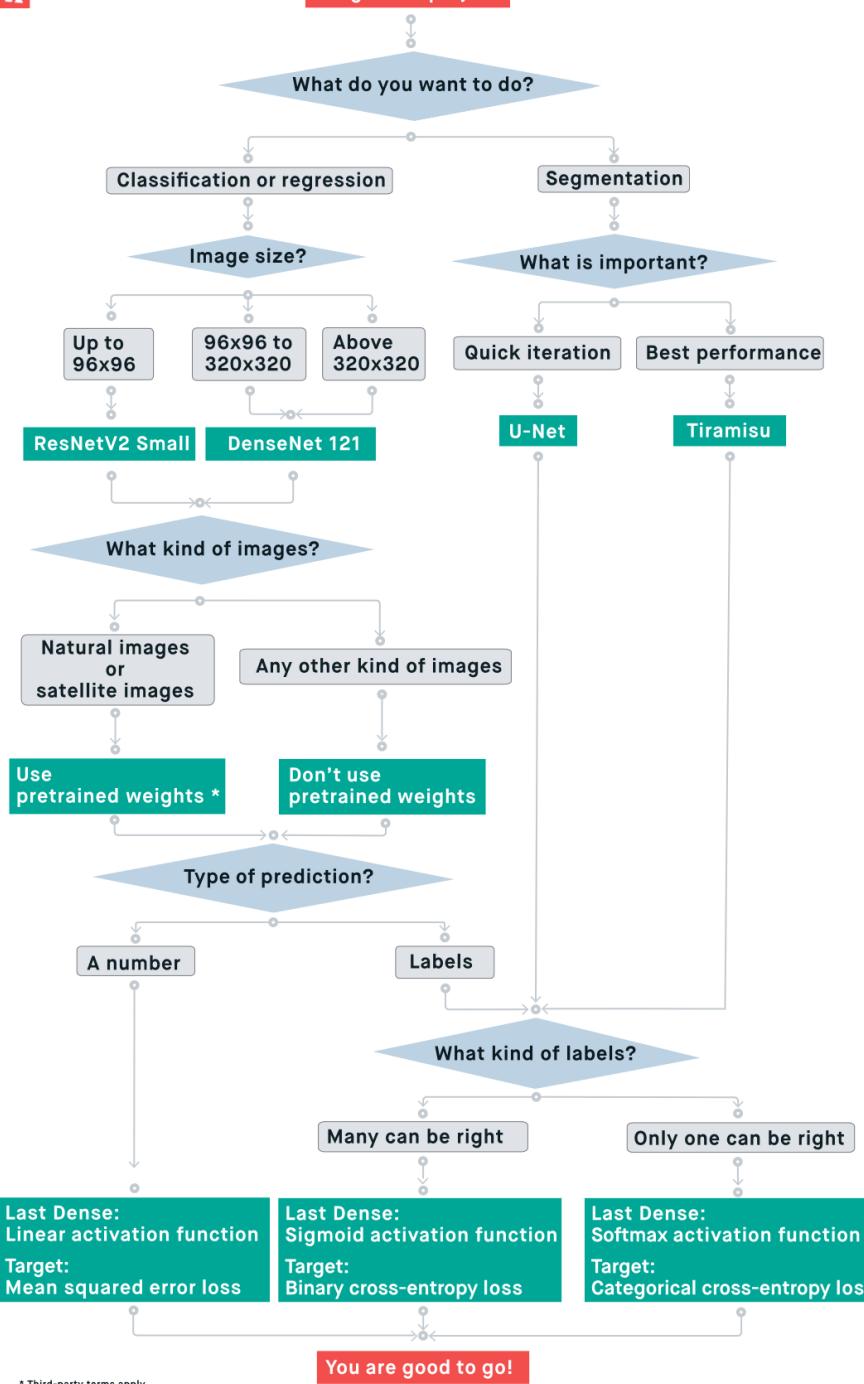
### Step 2:

- Do on your own



40 mins

**Individual Activity**



# Cheatsheet

<https://peltarion.com/knowledge-center/documentation/cheat-sheets>



# Linking Them Together

## App Development

### Top 9 No-Code Web App Development Tools that May Compete with Bubble

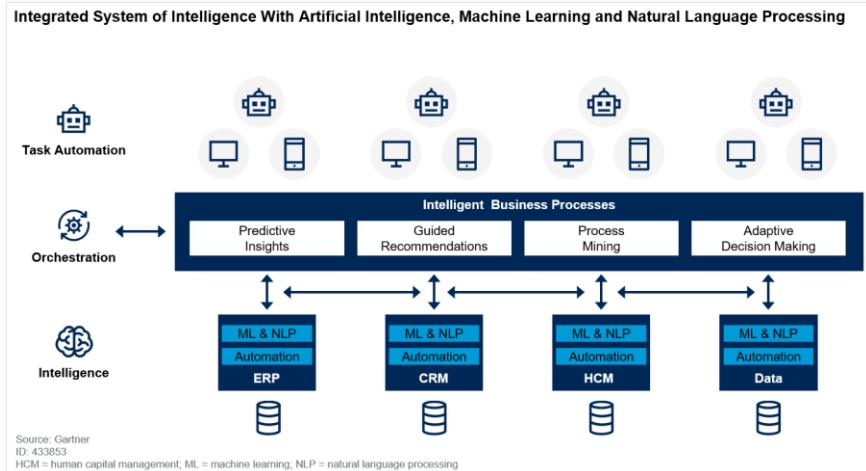
We're here to prove that "building the best product" is possible not only with Bubble.

Discover the 5 most powerful *Bubble* alternatives in the comparison table below to decide which one fits you best. Find more detailed information about the other *Bubble.io* alternatives after the table.

The screenshot shows the Bubble UI Builder interface. On the left is a sidebar with various tools like Design, Workflow, Data, Styles, Plugins, Settings, and Logs. The main workspace displays a 'Bubble Boilerplate' page with a 'GET STARTED' button. A modal window titled 'Button GET STARTED' is open, showing settings for appearance, conditional logic, and transitions. The properties panel on the right shows the element's width (110), height (228), and position (V: 180, H: 45, X: 0, Y: 0).

<https://uibakery.io/bubble-alternatives>

## RPA



Adobe Acrobat Document

<https://www.youtube.com/watch?v=FV8IM9SIFQ8> 56



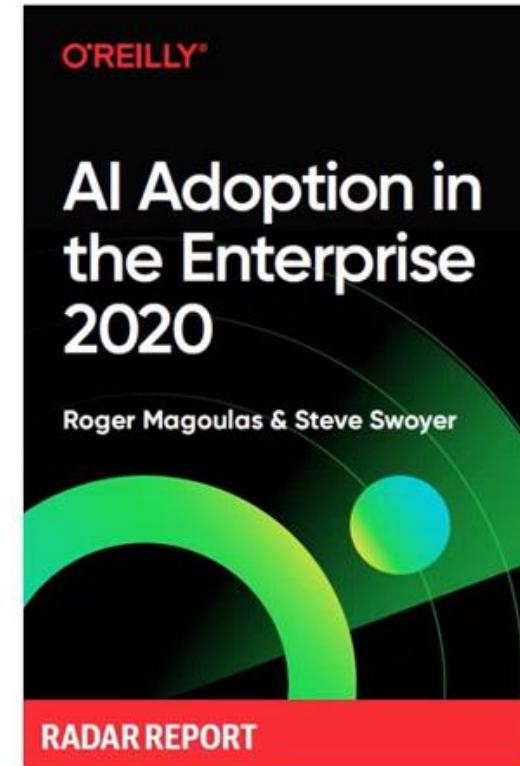
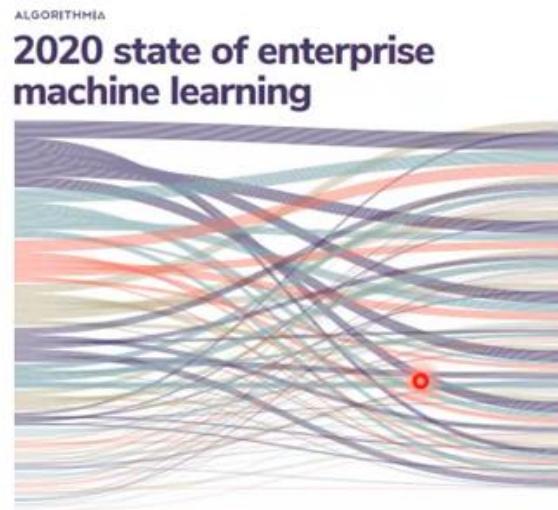
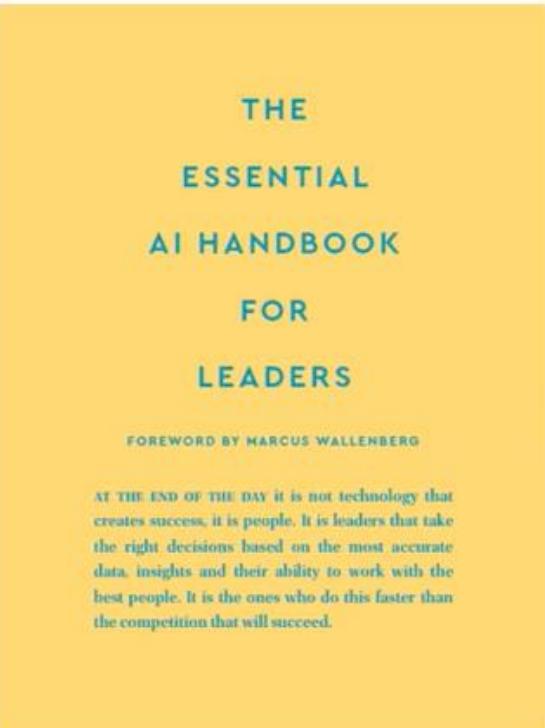
# When to use Machine Learning

---

- **What are our most pressing problems right now?**
  - Just like any other tool in business, AI should be viewed as a tool that can help make your organization more effective, profitable or streamlined
- **What parts of our business generate revenue but currently have low profit margins?**
  - These revenue streams could provide fertile ground for automation and acceleration via AI.
- **Where would we like to cut costs?**
  - Review your costs and pinpoint the ones you'd like to reduce. AI can help you better understand what generates costs and identify areas that could be optimized or changed to reduce them.
- **Where do we make a high percentage of errors in our work?**
  - A well-trained AI model has the capacity to perform with far less margin of error than humans
- **What work do our employees do that they don't particularly like?**
  - If it's repetitive or annoying for a human to do, there might be a component of the task better done by AI.



# Some easy readings





# Datasets and Data Prep

## GitHub

<https://kwseow.github.io/>

## kaggle

<https://www.kaggle.com/datasets>

Google  
Dataset Search Beta

<https://datasetsearch.research.google.com/>



<https://www.kdnuggets.com/datasets/index.html>



++ + a b | e a u

Microsoft



roboflow



# Dataset and Data Prep

---

- Datasets
  - <http://kwseow.github.io>
  - <https://datasetsearch.research.google.com/>
  - <https://www.kaggle.com/>
- Data prep
  - Excel
  - Tableau Prep
  - Power BI

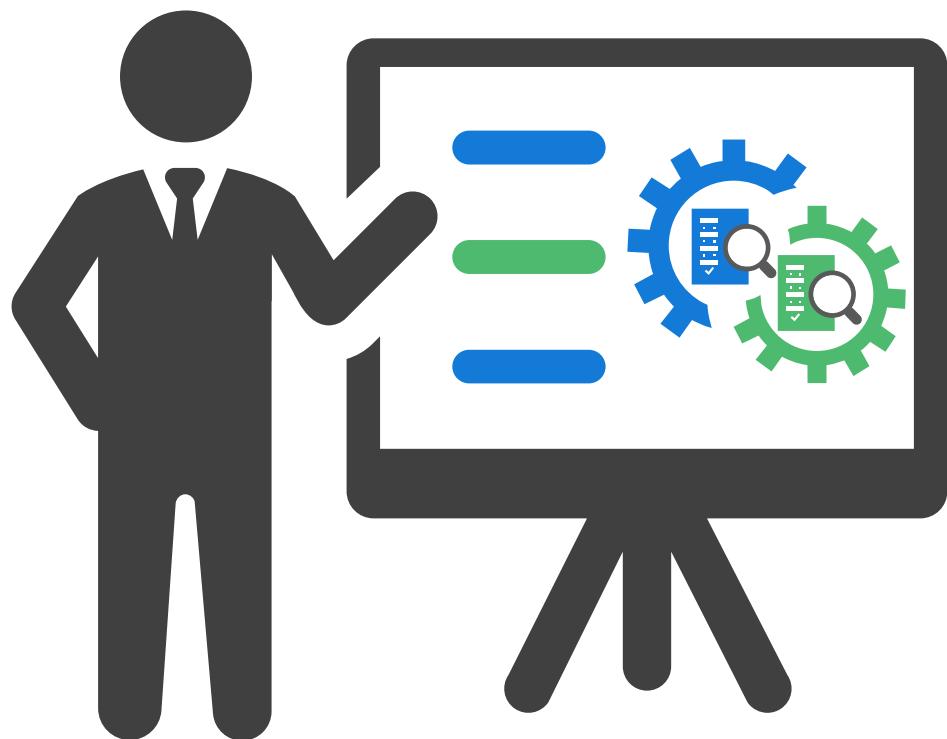
# Survey Time!

---

Please help to complete the survey before you leave the session. ^\_^  
(TRAQOM)



# Summary



Email  
[zack\\_toh@rp.edu.sg](mailto:zack_toh@rp.edu.sg)

Telegram  
[@zacktohsh](https://t.me/zacktohsh)

Source code:

62



# Thank you