



AI & MACHINE LEARNINGG MASTER CLASS

UTS::DataLounge

A/Prof Richard Xu



You hear about AI every day...

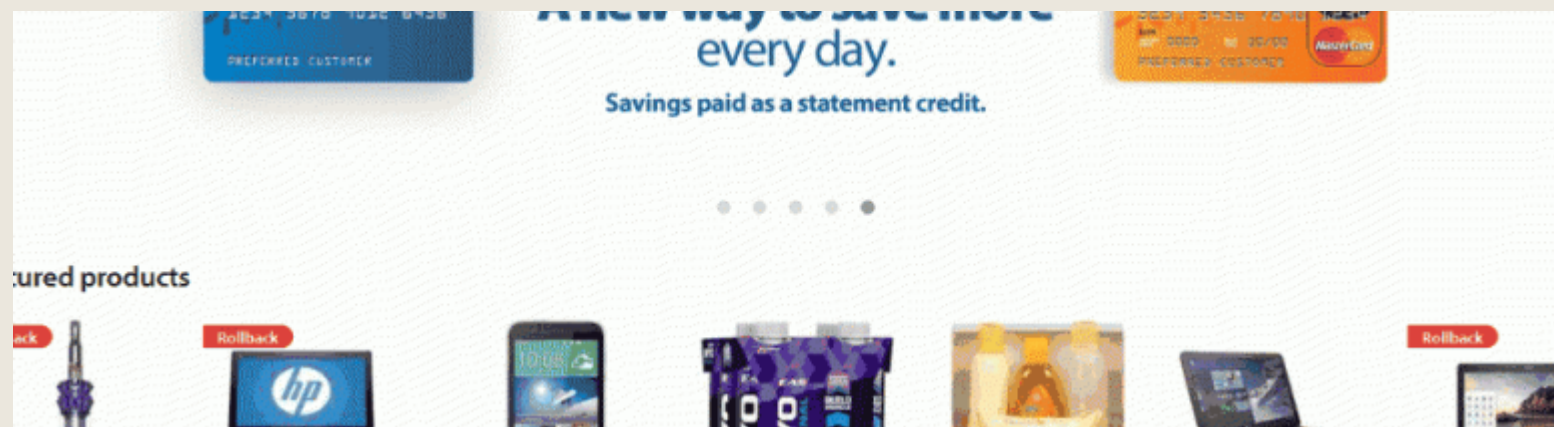


Machine Learning and AI is everywhere

- They are more commonly used than you think. Let's see three concrete examples
 - Retailer
 - Human Resource

[Retailer]

- A retailer has an **online shopping site**:



Data collected:

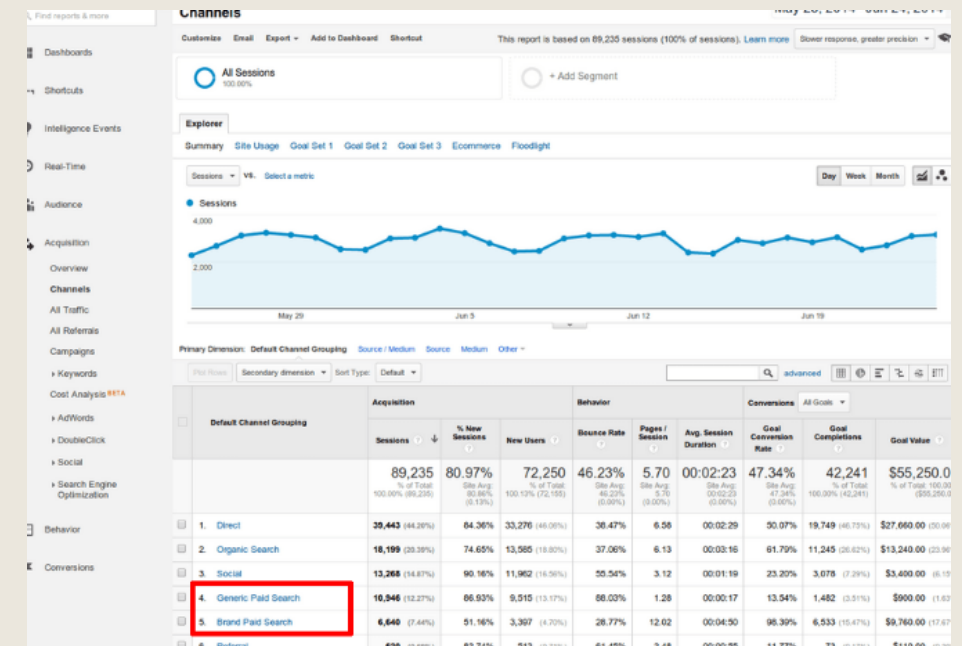
- Customer record
- Product Information
- Sales histories
- Royalty data
- ...

ML can help to:

- Sales prediction
- Customer segmentation
- Predictive inventory
- Construct a Chabot
- ...

[Retailer]

- A retailer may also use a customer tracking tool, such as **Google Analytics**:



Data collected:

- Customer browse information (time and pages browsed)
- Campaign response

ML can help to:

- Recommendation system
- Personalised email/campaign
- Product correlations
- Fine tuned recommendation system

[Retailer]

- A retailer may also have a physical store:



Data collected:

- Shopping basket
- Shopper trajectories

ML can help to:

- Optimised Store design
- Popular attraction spots

[Retailer]

- A retailer may also monitor its social account



Data collected:

- Tweeter messages about the retailer and its brands
- Facebook review on this retailer

ML can help to:

- Sentiment analysis about its customer experience
- Predicting trends

In summary, ML & AI helps a [Retailer] to:

- Understand customers better
- Individualised/targeted marketing *what promotion/product a customer may be most interested, personalised email*
- Create a positive experience *market driven store layout design*
- Better and prompt marketing conversations *using chatbot*
- Assist marketing professionals making decisions: *which customer segment should spend what amount*
-

[Human Resource]

- HR has a complete histories of each employee

ID	First Name	Last Name	Department	Track Leave As ...	Annual Leave - Annual Entitlement	Annual Leave - Accrued	Annual Leave - Carried from Last Year	Bonus Leave	Annual Leave - Total Entitlement Calculated	Start Date
1	Lucy	Pointer	Marketing	Days	21	0	3	0	24	08/Jan/2009
2	Tony	George	Admin	Days	20	0	2	1	23	07/Jan/2009
3	Jayne	Charles	Director	Days	23	0	0	0	23	09/Jan/2009
4	Mary	Mordon	IT	Days	20	0	0	3	23	01/Jan/2003
5	Tony	Lock	IT	Days	24	0	0	0	24	01/Jan/2003
6	Jack	Smith	Sales	Days	23	0	1	0	24	01/Jan/2003
7	Sarah	Hardey	Sales	Days	20	0	2	0	22	01/Apr/2005
8	Michael	Potts	IT	Hours	175	0	3	0	178	01/Jan/2003
9	Ian	Smith	Admin	Hours	175	0	0	0	175	01/Jan/2003
10										

Data collected:

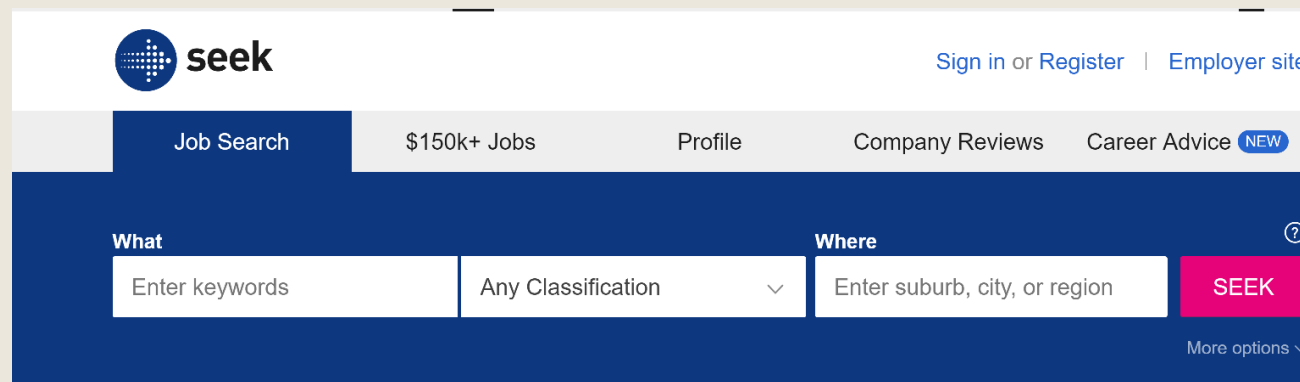
- Employee information
- Job and tasks information
- Department information
- Evaluation report
- ...

ML can help to:

- Performance indicator prediction
- Career planning
- Recommend trainings
- ...

[Human Resource]

- HR has source external job data

A screenshot of the Seek job search website. The header features the 'seek' logo on the left and links for 'Sign in or Register' and 'Employer site' on the right. Below the header is a navigation bar with 'Job Search' (highlighted), '\$150k+ Jobs', 'Profile', 'Company Reviews', and 'Career Advice' (with a 'NEW' badge). The main search area has a dark blue background. It includes a 'What' section with a text input 'Enter keywords' and a dropdown menu 'Any Classification'. A 'Where' section has a text input 'Enter suburb, city, or region'. A pink 'SEEK' button is to the right of the 'Where' input. A 'More options' link with a dropdown arrow is at the bottom right of the search area.

Data collected:

- External Job sites data
- ...

ML can help to:

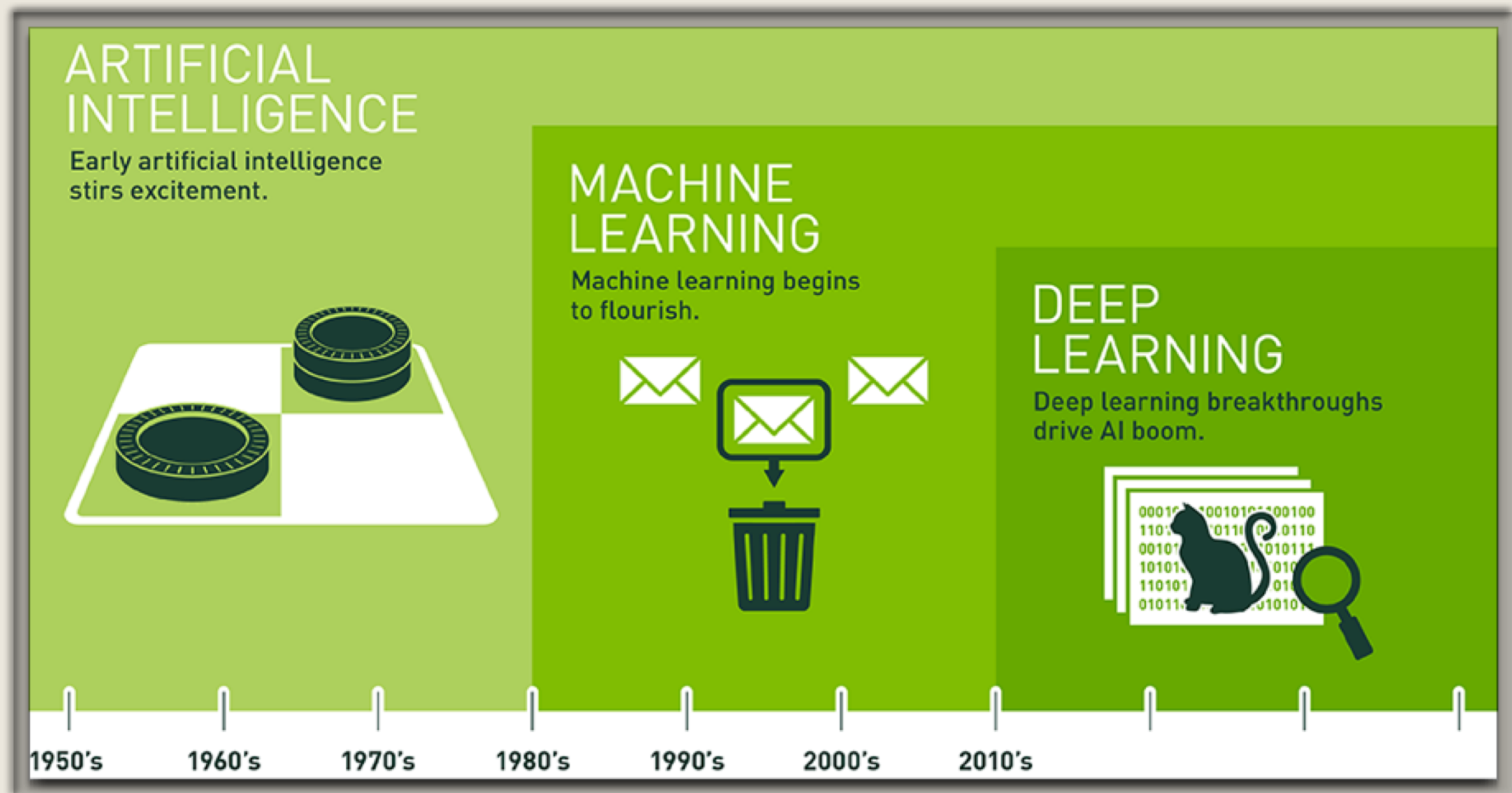
- predict employee's market value
- Retention prediction

This course

Not trying to make you an AI nor ML expert in 3.5 hours, but instead:

- Give you an overview of AI/Machine Learning (include Deep Learning)
- Some of the very latest technologies in AI (include our own research @ UTS)
- Fundamental techniques of ML: Regression, Classification and clustering algorithms
- Get hands dirty: show some line of Python code
- Where to go from here

So what are the buzzword: AI, ML & DL?



- a “biased” view from NVIDIA:
- but not far from the truth

Some latest state-of-the-art technologies in ML



- Detect cloth landmarks
- Useful for predicting fashion trend

Real-time 2D pose estimation



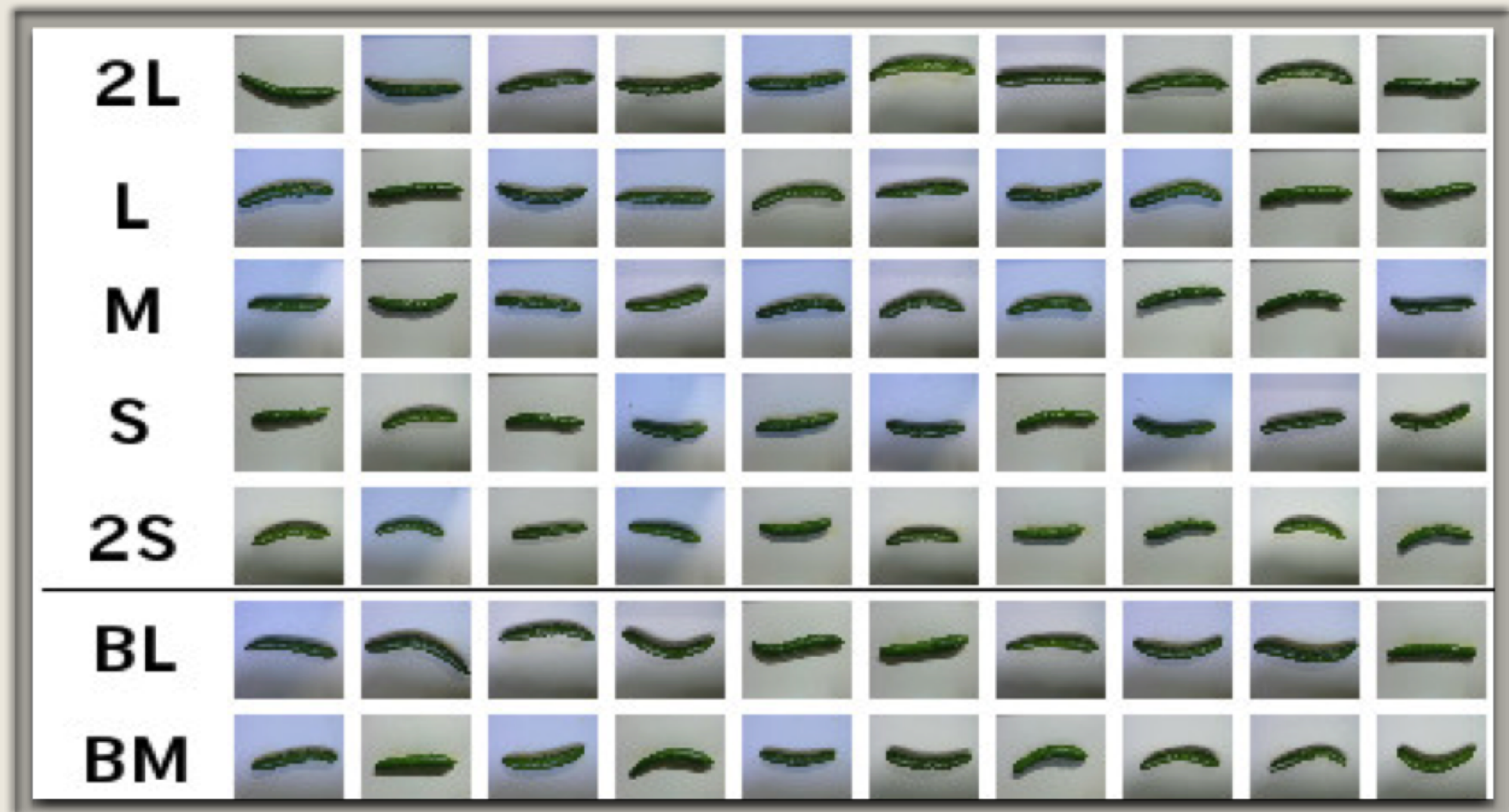
- Learning human pose: surveillance application
- and sports

Real-time 2D pose estimation



- Learn image labelling by parts

Cucumber sorting by Japanese farmer



Google sunroof



- Personalized solar savings estimator, powered by Google Earth imagery
- Same type of technologies for estimating housing price
- Can you think about another good application of house size detection on Google earth images?

UTS::Deep Learning

Automated Job Ad Generator

JD Auto Completion Tool

Classification

it / computing / software ↕

Number of words

5

Submit Form

Initial Text

a global brand and an industry leader that is based sydney urgently require a ux designer to join their team of highly talented and creative individuals this is a very high profile project that will be used nation wide . we are looking for

<http://www.ridl.com/jobgenerator/>

UTS::Deep Learning

Automated Sports action recognition



- Recognize 4 given action sequences in tennis

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

- a new version:

<https://www.youtube.com/channel/UChLPFBBeNIA0arPTrCor3tTg>

UTS::Deep Learning

Automated natural language processing to
SQL queries translation



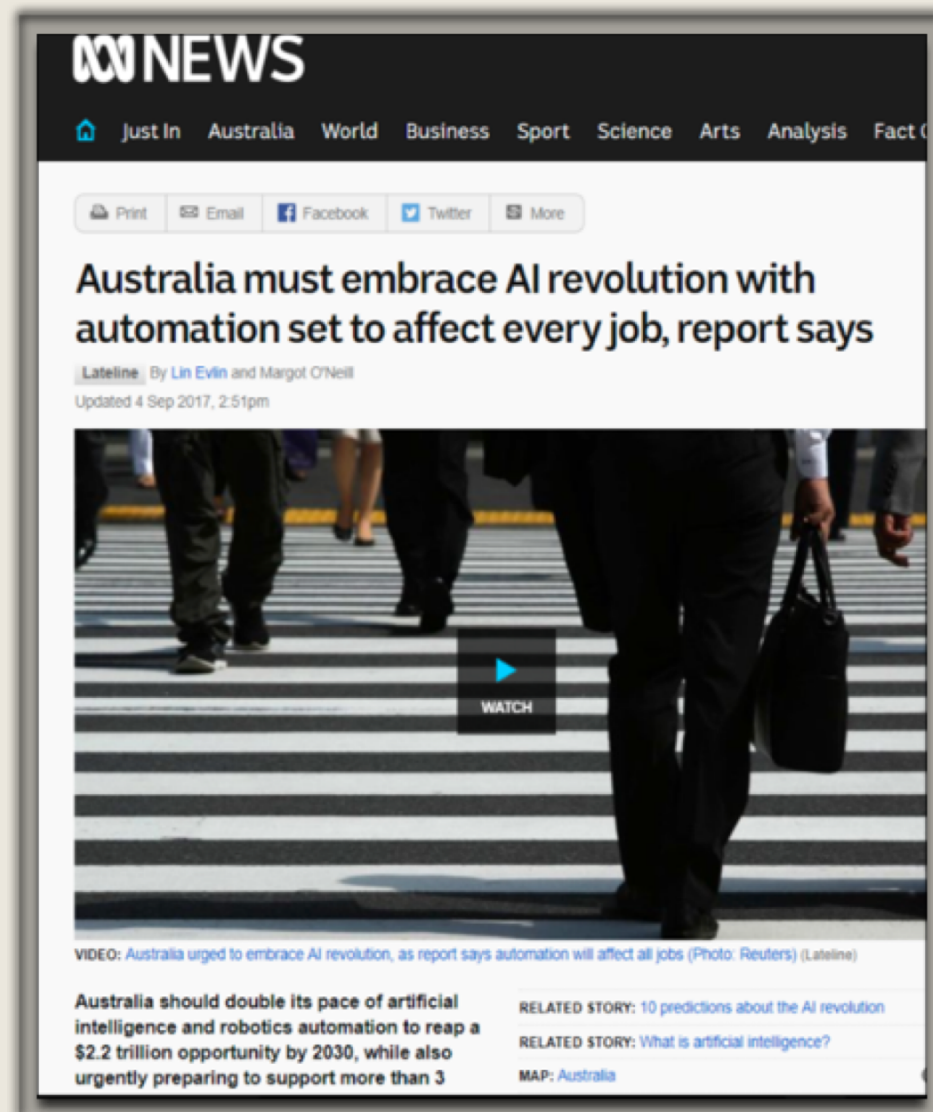
The impact AI/ML/DL will have on jobs in the future

Frey and Osborne 2014 report,
Probability of computerisation
of an occupation by AI

https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

Probability	Occupation
0.54	Advertising Sales Agents
0.63	Construction and Building Inspectors
0.61	Market Research Analysts and Marketing Specialists
0.71	Construction and Related Workers, All Other
0.75	Painters, Construction and Maintenance
0.85	Sales Representatives, Wholesale and Manufacturing, Except Technical and Scientific products
0.86	Real Estate Sales Agents
0.88	Construction Laborers
0.92	Retail Salespersons
0.94	Accountants and Auditors
0.94	Paralegals and Legal Assistants
0.98	Legal Secretaries
0.99	Telemarketers

Is Australia prepared? How about elsewhere?



- <http://www.abc.net.au/news/2017-08-08/australia-must-embrace-ai-revolution-alphabet-report/8774044>



- <https://www.technologyreview.com/s/609038/chinas-ai-awakening/>

Is this a good news for machine learning employees?

DAILY NEWS[VIEW ALL NEWS](#)

AUSTRALIA – NUMBER OF AI RELATED JOBS ON THE RISE, INDEED FINDS

28 February 2018

[Email](#) [G+](#) [f](#) [t](#) [in](#)

The number of AI related jobs listed on Indeed Australia has doubled since 2015, according to an analysis from the job board.

Job ads which referred to 'artificial intelligence' and 'machine learning' in the job description increased by 50% in January when compared January 2017. The most common job title with references to AI was

**InnovationAus.com**
Public Policy and Business Innovation

[Home](#) [About](#) [News](#) [Events](#) [Partners](#)

[AIMEE CHANTHADAVONG](#) | MARCH 6, 2018

Queensland's AI skills drought

[Skills](#)



[SIGN UP](#)
Your email

technology companies, and Melbourne produces more ...

Artificial intelligence is accelerating — investors need to keep up
[The Australian](#) - 54 minutes ago

In 2015 an open letter, signed by Bill Gates, Elon Musk, Stephen Hawking and by industry experts such as Demis Hassabis, Shane Legg and Mustafa Suleyman, the co-founders of British **AI company** DeepMind, now owned by Google's parent Alphabet, "continuous control" over **artificial intelligence** was ...

 **The Ai Group is calling for an international research organisation in ...**
[The Advertiser](#) - 7 Mar. 2018

- Clearly, there is a skill shortage globally and in Australia

Moral of the story

- It's never been a better time to study AI, Machine learning and Deep learning!

Also it's Time for some needy greedy technical stuff!

Supervised VS Unsupervised Learning

- **Supervised** – There is a label (human supervision) associated with every data
- **Unsupervised** – there is no label associated with the dataset





SUPERVISED LEARNING

Supervised Learning

Input	Output
1	2
2	3
4	5
6	7
12	13
15	?

- You have input/output pairs, and lots of them! We call them training data
- Now there is a new input, what may its output be?

Supervised Learning

Input	Model	Parameter	Output
1	+	1	2
2			3
4			5
6			7
12			13
15			?

- We assume there is a model
- The model should have a parameter that maps each input to its corresponding output (model)
- This model and parameter is **shared** amongst all data (inference)

Supervised Learning

Input	Model	Parameter	Output
1	+	1	1
2			3
4			5
6			7
12			13
15			?

- Here is key: once we know the model and parameter, we don't need training data any more to predict output for new input

In reality...

Input	Model	Parameter	Output
1.2	+	?	1.5
2.6			3.3
4.1			5.3
6.7			7.6
10.1			18.2
12.2			13.3

- Models are never easily guessed
- guessing the right model is the **key!**
- Also, no model and parameters can give perfect output for every input
- **The key take way:** given a model, we need to find a parameter that generates least amount of *total discrepancies* between **predicted output** and **true output**

Also in reality ...

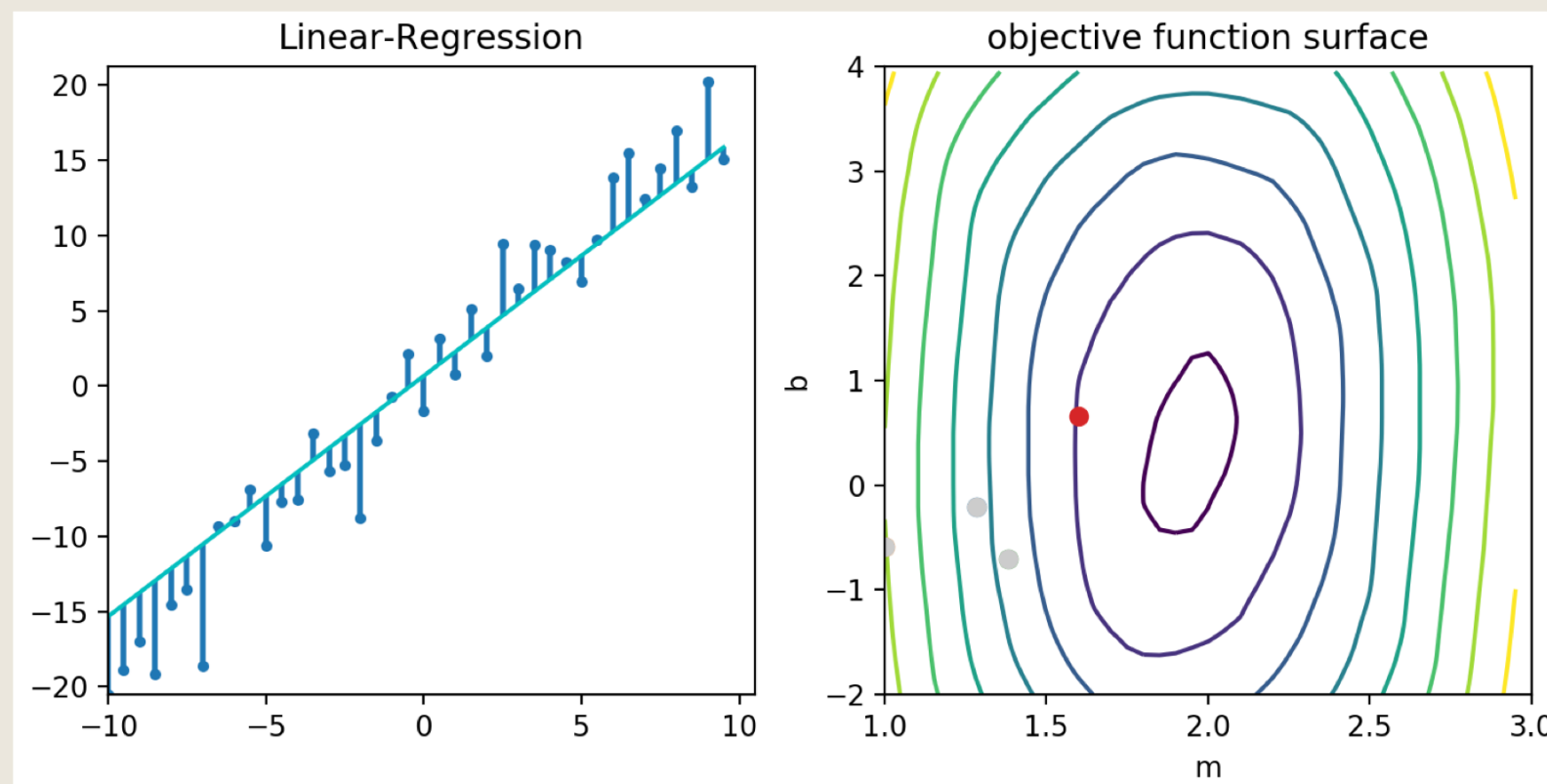
Input		Model	Parameter	Output
\$ shoes	\$ cloth	?	?	\$ cosmetics
264.5	2231			345
324.1	2342			531
326.7	983.6			760
343.1	417.5			182
432.2	192.6			133.6

- Input may be an array of numbers (vector)
- Output can also be (usually not) vector
- Guess the mode is difficult
- Once the model is fixed, parameter can be worked out **mathematically**

Parameters

- Every parameter gives a somewhat different overall discrepancy between y and *predicted* y
- See demo:

https://github.com/roboticcam/machine-learning-notes/blob/master/industry_master_class.ipynb



Also in reality even harder...

Input	Model	Parameter	Output
	Sophisticated	Complicated and high dimensional	"Dog"
			"Cat"
			"Fish"

- Input may complex such as text, image, graph ...
- Need sophisticated Model
- Parameters are hard to obtain, and require clever **mathematics**

Supervised Learning - Regression or Classification

- **Regression** - Predict results within a **continuous output** e.g. age, house prices, stock prices, spending value
- **Classification** - Predict results in a **discrete output** E.g.
 - the type of animal an image belong to
 - whether a customer is male or female

[Retailer] Regression example

- Predict profitability of each customer

Input:

- Age
- Gender
- Household count
- Income range

Output:

Profit per customer of running a marketing campaign

[Retailer] Regression example

- Predict ice cream sales volume

Input:

- Base price
- Time of day
- Weather
- Loyalty card discounts

Output:

total ice cream sales

[HR] Regression example

- Produce recommend salary increase

Input

- number of years of service
- qualification score
- Job type Market value
- Performance review score

Output

- Salary increase

[Building] Regression Example

Input:

- land size (m²)
- Floor size (m²)
- Number of bedrooms
- Number of bathrooms
- Number of car space
- standardized sale price of the suburb
- How old is the house

Output:

- House's sale price

[Retailer] Classification example

- Predict if a customer will be interested in the new product (hence to decide if he/she should be included in the marketing campaign)

Input:

- Age
- Gender
- Household count
- Income range

Output:

interested (**1**) or not interested(**0**)

[Retailer] Classification example

- recommend if a brand of ice-cream be included on the shelf

Input:

- Base price
- Weather
- Sales history of this brand of ice-cream
- Average income of population around the physical store

Output:

On-shelf (**1**) or not on-shelf (**0**)

[HR] Classification example

- Predict if an employee may resign in the next six months

Input

- number of years of service
- qualification score
- Job type Market value
- Performance review score
- Current pay
- How recent was the last LinkedIn profile change

Output

- Yes (**1**) or No (**0**)

[Building] Classification

Example

Predict if a house will be sold in the next 6 weeks

Input:

- The asking price
- Time on the market
- land size (m²)
- Floor size (m²)
- Number of bedrooms
- Number of bathrooms
- Number of car space
- standardized sale price of the suburb
- How old is the house

- **Output:**
Sale (1) or No-sale (0)

Under many settings

- Same input
- but different outputs
- ✓ one correspond to regression
- ✓ another correspond to classification

Regression and classification

Energy consumption/efficiency of a building

Input:

- **Building area**
- **Window U-value**
- **Wall U-value**
- ...

Output (Regression)

How much energy is the building consuming? E.g. 1500kW, 3000kW

Output (Classification)

Is the building energy efficient?
(Y/N)

Regression and classification

Energy consumption/efficiency of a building

Input:

- **every subject marks**
- **Research time spent**
- **intern experience**

Output (Regression)

predicted grad salary

Output (Classification)

Will study honours? (Y/N)

Data types

■ **Question:** What is the difference between these two categorical output:

- User ratings of online products: 1,2,3,4,5
- Email domains:

1 = @gmail.com, 2 = @uts.edu.au, 3 = @hotmail.com,
4 = @amazon.com, 5 = @yahoo.com

Crowdsourcing: free labelling

- Labelling data is too expensive
- Can we get it done cheaply?



- Question is What if the majority of people input random answers?
Will it still work?

Supervised learning Models

Many models can be applied:

- For **regression**: Linear, Polynomial, Ridge, Lasso Regression, ElasticNet, Gaussian Process , Decision Tree . . . many more
- For **classification**: Neural networks, Support Vector Machine, Multinomial Logistic Regression (Softmax), Decision Tree, Random Forest . . . many more
- For details, check out my machine learning notes!

How can you tell if your classifier is accurate?

- Split your training data into “training” and “testing”

Input		Model	Parameter	Output
\$ mid-term	\$ assignment	?	?	\$ pass exam?
74	13			N
56	54			Y
58	91			Y
92	81			Y
43	86			N
81	91			Y
46	35			N

- Split your training data into “training” and “testing”
- Should we repeat this multiple times?

Classification accuracy

- Classification accuracy: Percentage of records correctly identified
- The following shows accuracy of 6/8

Input		True output	Predicted output
mid-term	assignment	pass exam	pass exam
74	13	N	N
56	54	Y	N
58	91	Y	Y
92	81	Y	Y
43	86	N	Y
81	91	Y	Y
46	35	N	N

Classification accuracy: skewed data

- What happens when the data set is skewed, i.e., there is a **class imbalance**?
- For example, a retailer website plans to roll out an expensive campaign for any shopper whom may spend more than \$2,000 in the next season. To do this, it requires the ML expert to develop a model to predict which shopper is in this category.
- From historical dataset, it is known that there is only 0.1% of all shoppers fall into this category.
- What if **classification accuracy** is used under this setting?

TP, TN, FP, FN

- **TP (True positive):** Classifier shows it is positive, it is really positive
- **TN (True negative):** Classifier shows it is negative, it is really negative
- **FP (False positive):** Classifier shows it is positive, it is really negative
- **FN (False negative):** Classifier shows it is negative, it is really positive

Precision, Recall and F1 score

- Precision or positive predicted value: $\frac{TP}{TP + FP}$

- Recall or sensitivity: $\frac{TP}{TP + FN}$

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$



UN-SUPERVISED LEARNING



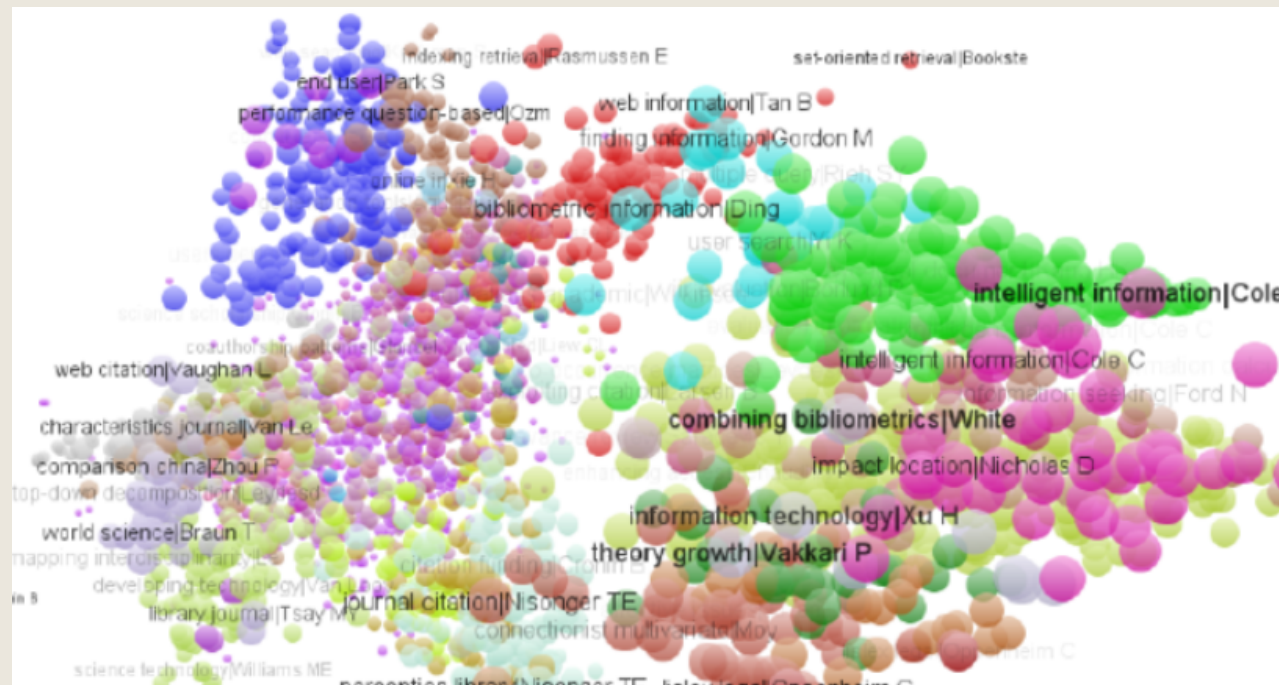
Unsupervised learning

- Not all dataset came with a labels/output
- For many dataset, the labels may not even necessary having a straight forward meaning
 - For example: in customer segmentation – what is the meaning for each customer segment?

Some models in unsupervised learning

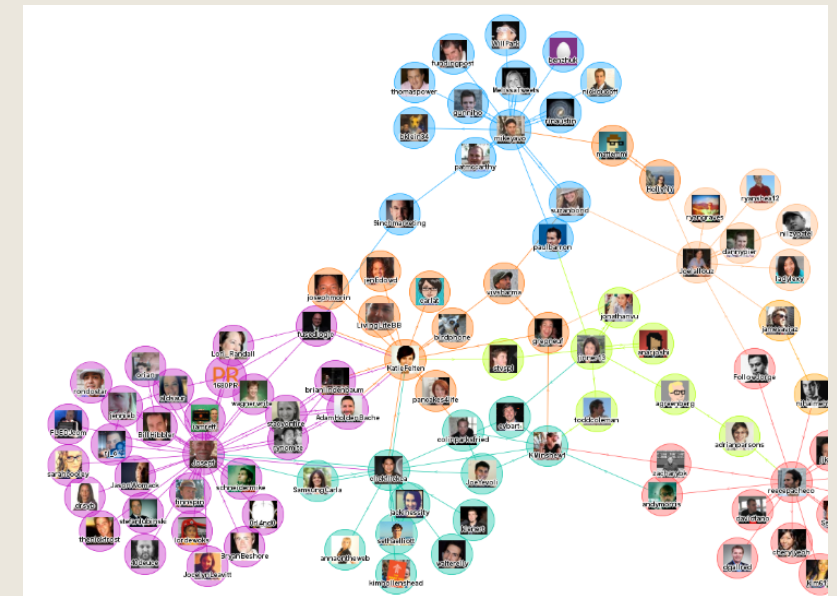
- Clustering k-means and mixture models
- Latent Dirichlet Allocation (**has been hot**)
- Autoencoders
- Generative Adversarial Networks (**very hot!**)

Application of Clustering



Document clustering

Image credit: ResearchGate



Social networks

Image credit: TouchGraph.com

Customer segmentation
Image credit: Smart insights



Simplest Clustering algorithm: K-means

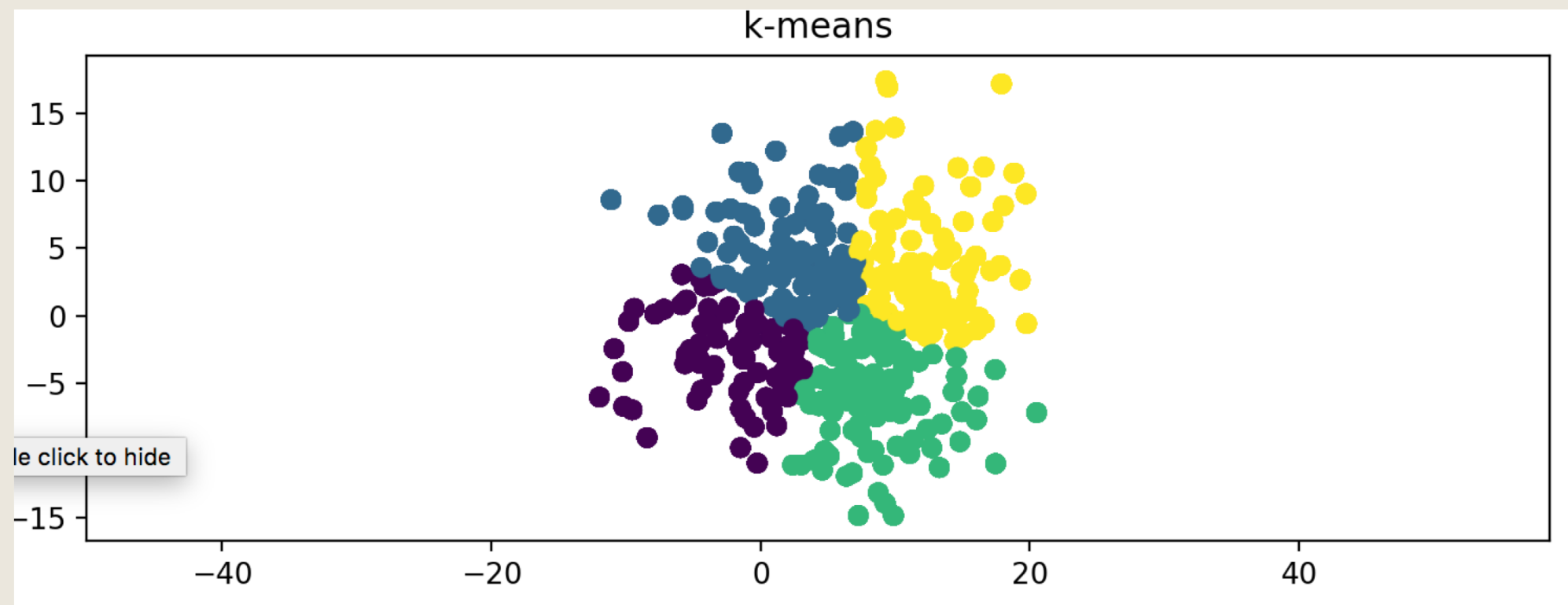
- **STEP 1:** Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
- **STEP 2:** Assign each object to the group that has the closest centroid
- **STEP 3:** When all objects have been assigned, recalculate the positions of the K centroids.

Repeat Steps 2 and 3 until the centroids no longer move.

K-means demo

- See demo:

https://github.com/roboticcam/machine-learning-notes/blob/master/industry_master_class.ipynb



Limitation of K-means

- assumes variance of the distribution of each attribute is spherical,
- all variables have the same variance;
- probability for all k clusters is the same, i.e., each cluster has roughly equal number of observations
- **Solution:** using a (Gaussian) Mixture Model

Exercise on Supervised (regression and classification)/Unsupervised

- Looking at Microsoft <https://how-old.net/> it was trained using images of people with known ages
- Obtain **segments** of passenger traveling behaviours from Transport Survey Data
- Given a video sequence: separate any arbitrary (moving) foreground from its background assuming background pixels varies less than the foreground (we compute statistics of the background first)
- Having a historical relationship between material consumption and labour hours, develop model to predict the labour hours for the material consumption in the future.
- Develop world's best image-based object recognition system
- Looking at Experian Mozaic <http://www.experian.com.au>, which train and categorize customers (with some data feature extraction) into a set of predefined categories.



MACHINE LEARNING PROJECT



Three layer approach: Layer 1: Application

- Business Analyst plays a significant role in this layer
- BA defines business' problem, obtain business value
- Domain specific knowledge is essential
- Each project is different: a BA will be needed at the initial phase for each individual project.
- They generally have a broad understanding of machine learning, and its capabilities.

Three layer approach: Layer 2: Model Formulation

- Data Scientist plays a main role in here.
- Transforms the business problem into a mathematical framework.
- This is usually the task of a data scientist who has knowledge on applying machine learning.

Three layer approach: Layer 3: Solver

- Machine Learning practitioner/researcher plays a major role in here
- Now that we have a model, we can solve these equations.
- Often, a generic solver may not be appropriate, especially when the data is BIG.
- This is why they need to have research level knowledge on machine learning and mathematics to make the solutions workable.

Traveller's tap-on time modelling

■ Layer 1: Application

From OPAL data, the business wants an estimate on the probability of passenger taps on at central station at various times (e.g., what is the probability of someone taps on at central at 8:15am?)

■ Layer 2: Model Formulation

Model all passenger tap on times using Bi-modal Gaussian Mixture Model (GMM)

■ Layer 3: Solver

Solve GMM using expectation-maximization;

see [one_d_opal_simulated.m](#)

Hotel recommendation system

■ Layer 1: Application

an online hotel business has a database containing every user's rating of their stayed hotels the marketing team wants to know which hotels to recommend to individual user in a promotional email (customized emails)

■ Layer 2: Model Formulation

data scientists decide to build a recommendation system using Non-Negative Matrix Factorization (NNMF) algorithm

■ Layer 3: Solver

there are many ways to solve NNMF, but the team decide to use Gradient Descend, because of the relative small size database

Learning analytics system

■ Layer 1: Application

hypothetically, UTS decides to build the world's best learning analytic system: it takes into consideration of the student's histories of studies and their interest, then it produces a "future study plan" deems to be best fit for each and every student and of course, they asked Richard's team to conduct this work

■ Layer 2: Model Formulation

Richard's team decides to base this model using a modified Recurrent Neural Network (RNN).

■ Layer 3: Solver

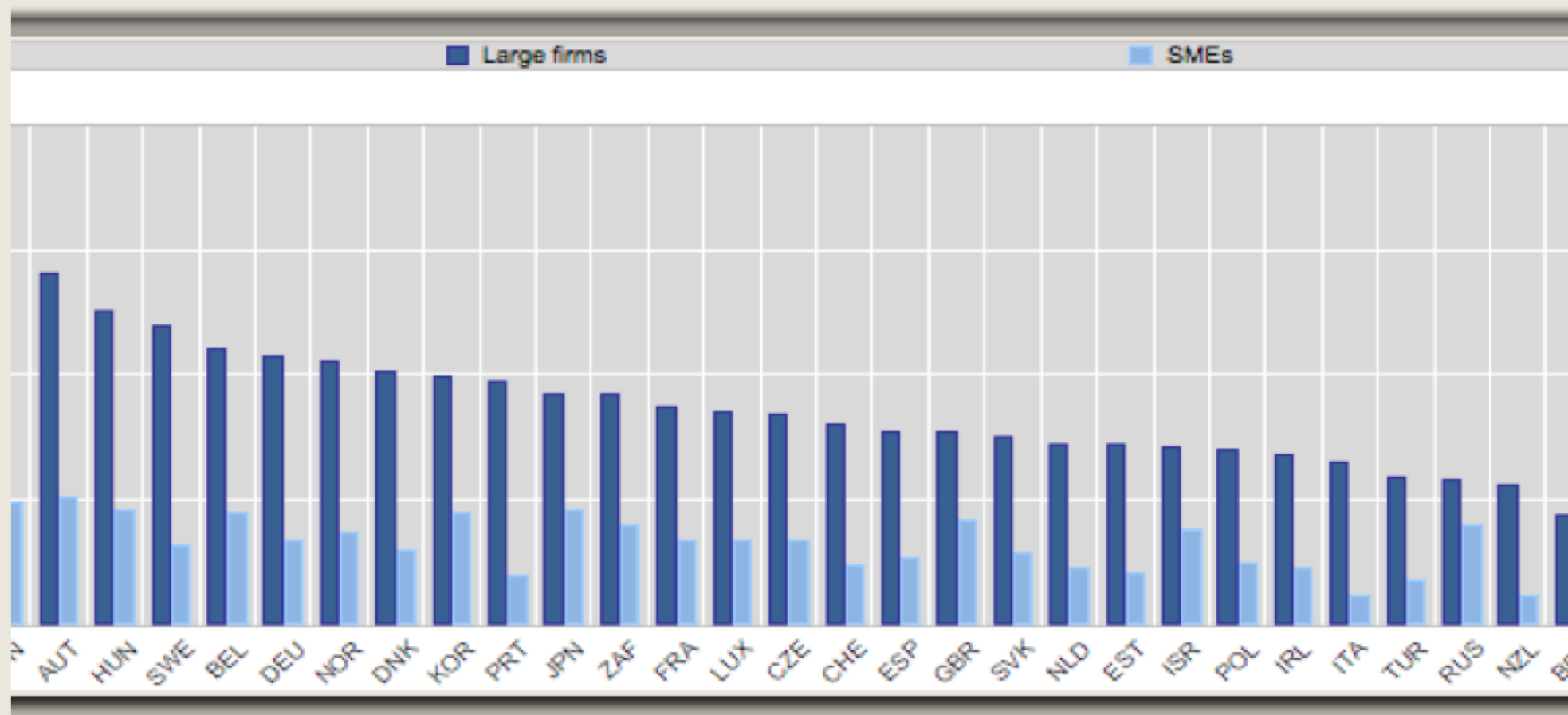
learn all the parameters of RNN using standard back-propagation.



SOME CONCLUSION REMARKS



OECD chart for collaboration between researches and businesses



OECD, based on Eurostat (CIS-2010) and national data sources, June 2013.

List of world leading AI and Machine Learning scholars employed by companies

- Geoffery Hinton, Google and DNNresearch Inc, USA
- Richard S. Sutton, head academic advisor to Royal Bank of Canada, Canada
- Michael I Jordan, technical adviser to the Ant Financials, China
- Yann Lecun, Director of AI Research at Facebook,USA
- Zoubin Ghahramani, Uber's Chief Scientist,USA
- Fei-Fei Li, Chief Scientist of Artificial Intelligence & Machine Learning, Google Cloud USA
- Daphne Koller, Chief Computing Officer, Calico Labs USA
- Andrea Frome, Director of Research, Clarifai UK
- Carol Reiley, Co-Founder & President, Drive.ai USA
- Hua Wu, Technical Chief, Baidu Natural Language Processing (NLP) Team China
- Ayse Naz Erkan, Staff Data Scientist, Twitter USA

shorten ML knowledge gap between industry and academia

- problem almost every data strategist is facing today is, how to make the most out of the booming opportunities in machine learning:
- when it comes to esoteric nature of data science, many corporate leaders maybe missing out on the full benefit, because they may be unsure of what machine learning is and what it can do for them
- Often, it can be as simple as fearing the unknown holding corporate leaders back from embarking on machine learning projects with specialists
- academia can help in many areas. For example, University can help industry to answer some of the questions (next slides)

“Is the product a really high tech one?”

- ML vendors tries to “sell” machine learning/AI solutions to industries every day.
- I how much technologies really is in its product?
- I for example, vendors may sell a “cutting-edge speech recognition” solution using a wrapper around off-the-shelf Google or Amazon cloud engine

“am I hiring the right data scientist?”

- produce a “machine learning demo” is much easier today than it has ever been
- there are many open-source stuff out there on Github! - simply run them will produce an “impressive demo”
- machine learning is new to a lot of data strategists, therefore how can a corporate pick the right candidate, i.e., someone knowledgeable in ML versus someone just ran the code from Github two days ago.

“am I using the ML tool appropriately?”

- if someone was to write a computer software, you know when something is wrong: for example, website crashes
- ML algorithm will “always” run and always produce some results
- therefore, it’s vitally important to make sure the results are meaningful and/or correct assumptions, algorithms are used

Some resources

- Prof Richard Xu's machine learning notes and source code

<https://github.com/roboticcam/machine-learning-notes>

- Andrew Ng's Machine learning MOOCS

<https://www.coursera.org/learn/machine-learning>