MONASH University

MONASH
INFORMATION
TECHNOLOGY

A peak into…
Artificial Intelligence (AI) and Machine Learning (ML)

Ian Wern Han Lim
lim.wern.han@monash.edu

Referencing materials by
Chern Hong LIM

GROUP OF EIGHT AUSTRALIA

MONASH University

Ready?

Agenda

- What is?
  - Artificial Intelligence
  - Machine learning

Agenda

- What is?
  - Artificial Intelligence
  - Machine learning
    - Supervised learning
    - Unsupervised learning
    - Semi-supervised
    - Reinforced learning
    - Deep learning

MONASH University

Let us begin…

About me…

- Wern Han LIM (Ian)

About me…

- Wern Han LIM (Ian)
  - B. Comp Science graduate 2009
  - PhD Comp Science graduate 2018

About me…

- Wern Han LIM (Ian)
  - B. Comp Science graduate 2009
  - PhD Comp Science graduate 2018

- Teaching since 2014
  - FIT1008 Introduction to Comp Science
  - FIT2004 Algorithms and Data Structures
  - FIT3155 Advanced Algorithms and Data Structures

  - FIT3152 Data Science/ Analytics
  - FIT2014 Theory of Computation
  - And a number of supervisions (FYP and Honours)

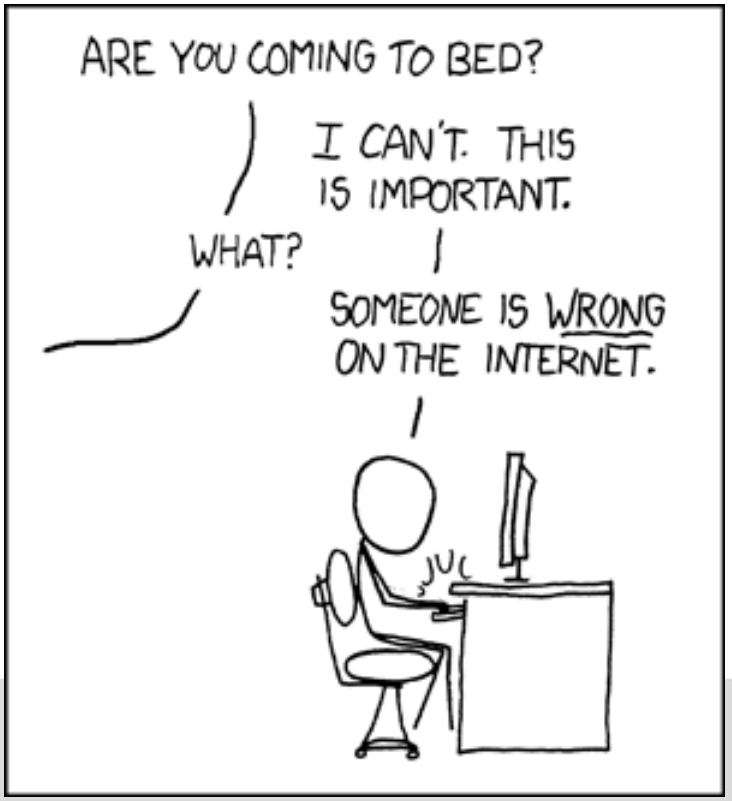
About me…

- My main area?
  - Information retrieval (IR)
  - User Generated Content (UGC)

About me…

- My main area?
  - Information retrieval (IR)
  - User Generated Content (UGC)

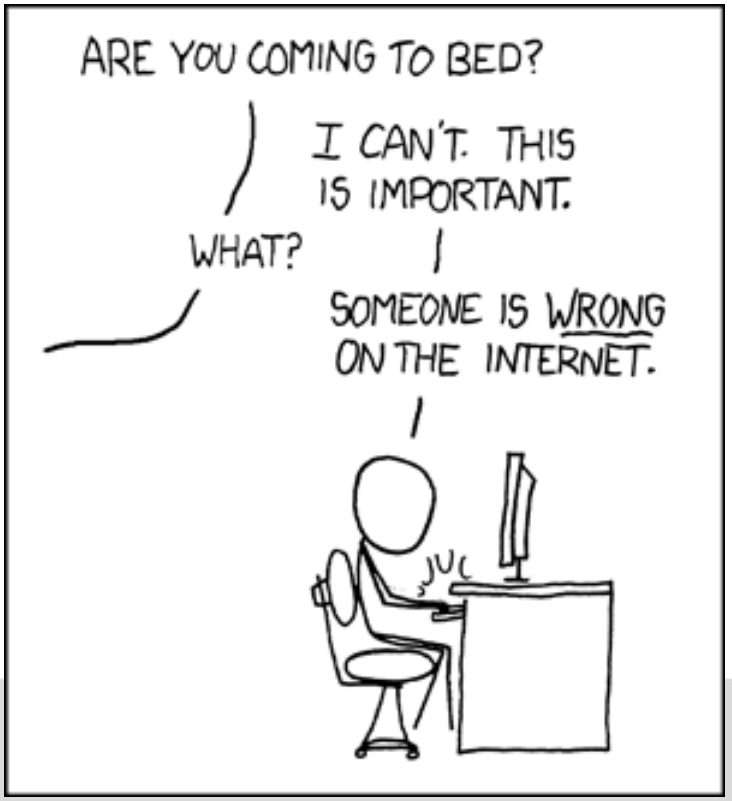
- My PhD Research?



About me…

- My main area?
  - Information retrieval (IR)
  - User Generated Content (UGC)

- My PhD Research?
  - Finding good #tags that describe documents, images etc.
  - Predict the best answer on QnA platforms (StackOverflow)
  - Organizing comments on Reddit
  - Profiling users on all 3 platforms
    - Expertise
    - Reliability



About me…

- I do have speech impairment
  - Pronunciation can be GG at times
    - Ask me to repeat
    - Ask me to spell
    - Ask me to write

About me…

- I do have speech impairment
  - Pronunciation can be GG at times
    - Ask me to repeat
    - Ask me to spell
    - Ask me to write
  - Feel free to interrupt me anytime
  - Keep it casual =)

MONASH University

Questions?

Artificial Intelligence (AI)

What is it?

MONASH University

Artificial Intelligence (AI)

What is it?

- Mimic the cognitive functions of human

Artificial Intelligence (AI)

What is it?

MONASH University

- Mimic the cognitive functions of human
  - Perception
  - Reasoning
  - Learning and knowledge organization

Artificial Intelligence (AI)

What is it?

MONASH University

- Mimic the cognitive functions of human
  - Perception
    - Natural language
    - Vision
  - Reasoning
  - Learning and knowledge organization

Artificial Intelligence (AI)

What is it?

MONASH University

- Mimic the cognitive functions of human
  - Perception
    - Natural language
    - Vision
  - Reasoning
    - Cause-and-effect
    - Xs affecting the Y
  - Learning and knowledge organization

Artificial Intelligence (AI)

What is it?

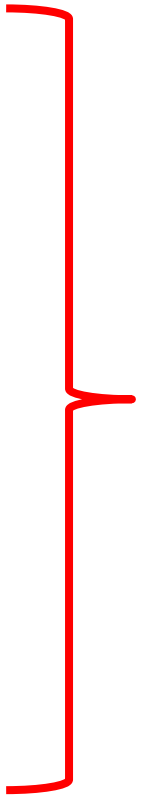
MONASH University

- Mimic the cognitive functions of human
  - Perception
    - Natural language
    - Vision
  - Reasoning
    - Cause-and-effect
    - Xs affecting the Y
  - Learning and knowledge organization
    - Using observations/ empirical data
    - Make a set of rules
    - Generalize what is learnt

Artificial Intelligence (AI)

What is it?

- Mimic the cognitive functions of human
  - Perception
    - Natural language
    - Vision
  - Reasoning
    - Cause-and-effect
    - Xs affecting the Y
  - Learning and knowledge organization
    - Using observations/ empirical data
    - Make a set of rules
    - Generalize what is learnt
    - Make the best decisions



General AI
(strong AI)

Artificial Intelligence (AI)

What is it?

Artificial Intelligence (AI)

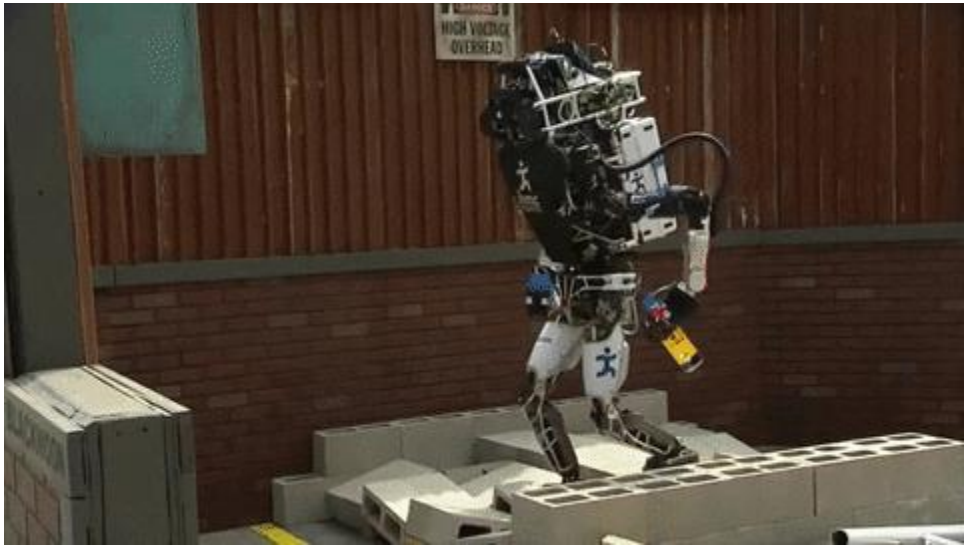
What is it?

"AI is a fundamental risk to the existence of human civilization."

- Elon Musk

Artificial Intelligence (AI)

What is it?



"AI is a fundamental risk to the existence of human civilization."

- Elon Musk

Artificial Intelligence (AI)

What is it?

"AI is a fundamental risk to the existence of human civilization."

- Elon Musk

Artificial Intelligence (AI)

What is it?

Artificial Intelligence (AI)

Dangerous?

- Thing is, we are still very far…

Artificial Intelligence (AI)

Dangerous?

- Thing is, we are still very far…

BUSINESS \ TECH \ ARTIFICIAL INTELLIGENCE

This AI startup claims to automate app making but actually just uses humans

11

Who could have seen that coming?

By Nick Statt | @nickstatt | Aug 14, 2019, 1:58pm EDT

Artificial Intelligence (AI)

Dangerous?

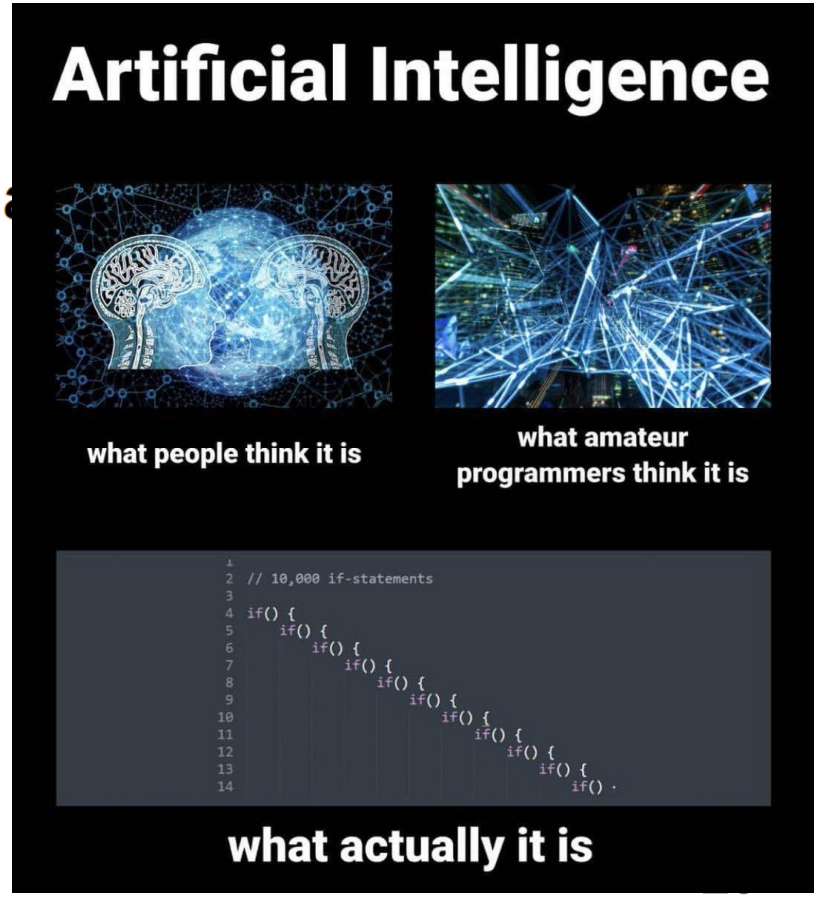
- Thing is, we are still very far…

BUSINESS \ TECH \ ARTIFICIAL INTELLIGENCE

This AI startup claims to automa… actually just uses humans

Who could have seen that coming?

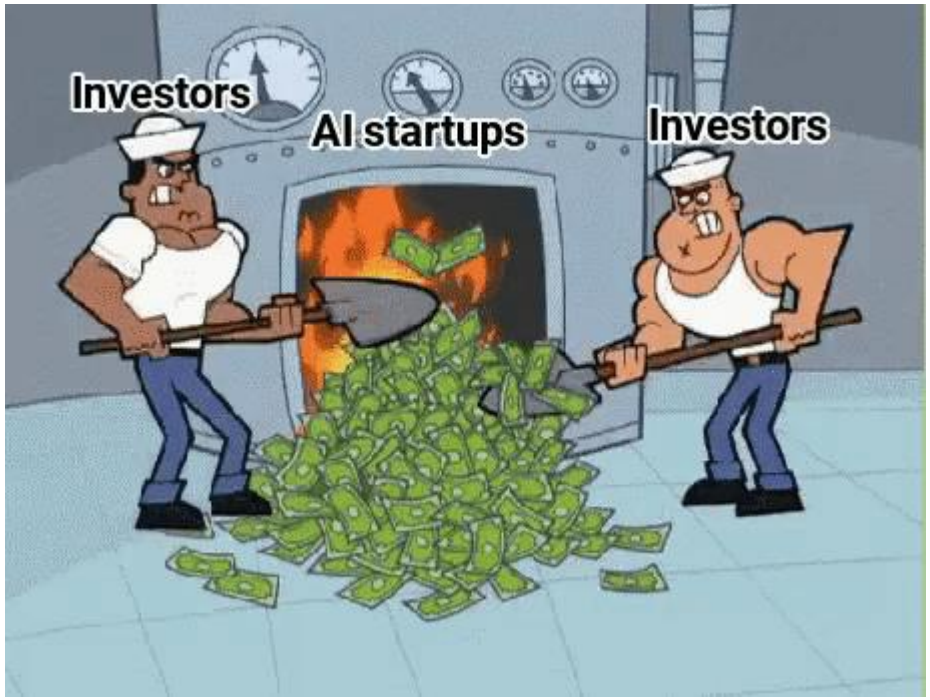
By Nick Statt | @nickstatt | Aug 14, 2019, 1:58pm EDT



Artificial Intelligence (AI)

Dangerous?

- Thing is, we are still very far…



Artificial Intelligence (AI)

Dangerous?

- Thing is, we are still very far…

- AI can do many things better than us
- AI can't do many things better than us

Artificial Intelligence (AI)

Dangerous?

- Thing is, we are still very far…

- AI can do many things better than us
  - Mundane task
  - Identify patterns
  - Anomaly detection
- AI can't do many things better than us

Artificial Intelligence (AI)

Dangerous?

- Thing is, we are still very far…

- AI can do many things better than us
  - Mundane task
  - Identify patterns
  - Anomaly detection
- AI can't do many things better than us
  - Make decision with a lot of variables
  - Take risks

Artificial Intelligence (AI)

Dangerous?

- Thing is, we are still very far…

- AI can do many things better than us
  - Mundane task
  - Identify patterns
  - Anomaly detection
  - … and many more
- AI can't do many things better than us
  - Make decision with a lot of variables
  - Take risks
  - … and many more

MONASH University

Questions?

Machine Learning (ML)

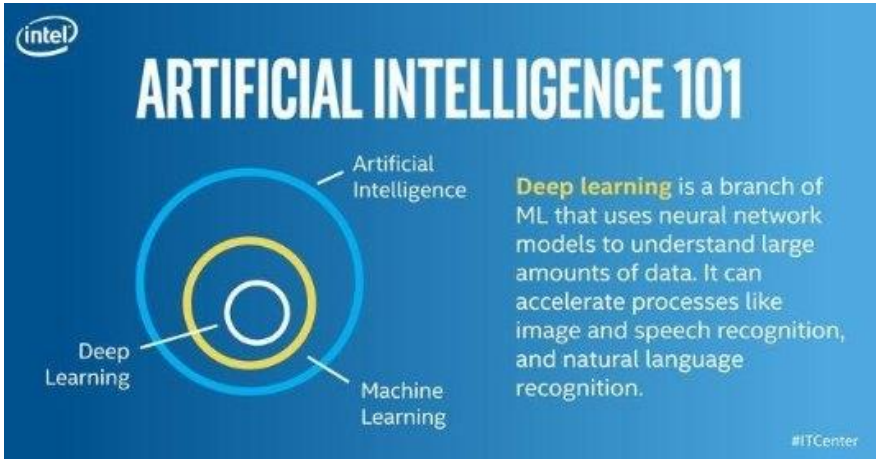
What is it?

MONASH University

Machine Learning (ML)

What is it?

- Subset of AI



Machine Learning (ML)

What is it?

MONASH University

- Subset of AI
  - Focus on the brain of the machine
  - Enable machine to evolve based on empirical data

Machine Learning (ML)

What is it?

MONASH University

- Subset of AI
  - Focus on the brain of the machine
  - Enable machine to evolve based on empirical data

- Use data to gain knowledge

Machine Learning (ML)

What is it?

MONASH University

- Subset of AI
  - Focus on the brain of the machine
  - Enable machine to evolve based on empirical data

- Use data to gain knowledge
  - More data = better

Machine Learning (ML)

What is it?

MONASH University

- Subset of AI
  - Focus on the brain of the machine
  - Enable machine to evolve based on empirical data

- Use data to gain knowledge
  - More data = better
  - Especially today with big data
    - Volume
    - Variety
    - Velocity

Machine Learning (ML)

What is it?

MONASH University

- Subset of AI
  - Focus on the brain of the machine
  - Enable machine to evolve based on empirical data

- Use data to gain knowledge
  - More data = better
  - Especially today with big data
    - Volume
    - Variety
    - Velocity
  - Quality do matter
    - “Drowning in the information and starving for knowledge” – John Nisbitt

Machine Learning (ML)

What is it?

MONASH University

- Subset of AI
  - Focus on the brain of the machine
  - Enable machine to evolve based on empirical data

- Use data to gain knowledge
  - More data = better
  - Especially today with big data
    - Volume
    - Variety
    - Velocity
  - Quality do matter
    - “Drowning in the information and starving for knowledge” – John Nisbitt

Machine Learning (ML)

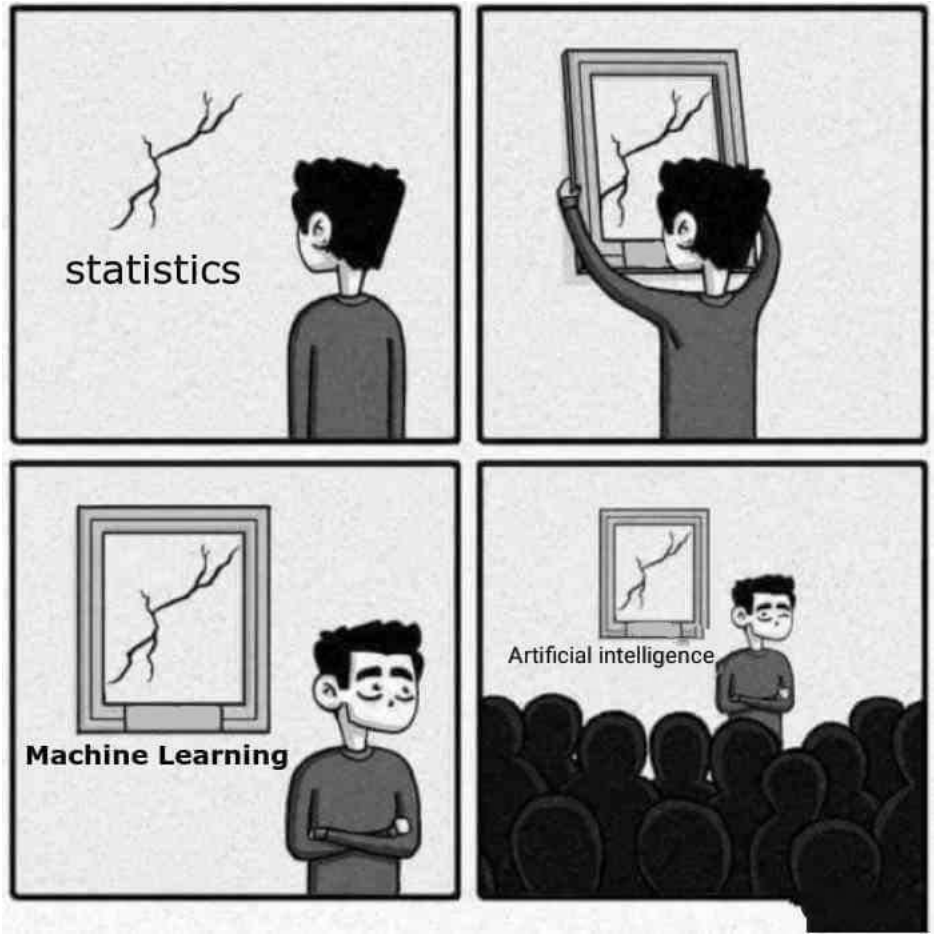
What is it?

MONASH University

- Subset of AI
  - Focus on the brain of the machine
  - Enable machine to evolve based on empirical data
- Use data to gain knowledge
  - More data = better
  - Especially today with big data
    - Volume
    - Variety
    - Velocity
  - Quality do matter
    - "Drowning in the information and starving for knowledge" – John Nisbitt

Machine Learning (ML)

What is it?



MONASH University

Questions?

Machine Learning (ML)

Types?

Machine Learning (ML)

Types?

- Supervised learning
- Unsupervised learning
- Semi-supervised learning
- Reinforced learning
- Deep learning?

MONASH University

Questions?

Supervised Learning
What is it?

- Given a set of observations

Supervised Learning
What is it?

MONASH University

- Given a set of observations
  - Features/ variables
  - Values/ measures

Supervised Learning
What is it?

MONASH University

- Given a set of observations
  - Features/ variables
  - Values/ measures
- … that is labelled
  - Predefined classes
  - Predefined values

Supervised Learning
What is it?

MONASH University

- Given a set of observations
  - Features/ variables
  - Values/ measures
- … that is labelled
  - Predefined classes
  - Predefined values

- Are we able to map the observations to their label?

Supervised Learning
What is it?

MONASH University

- Given a set of observations
  - Features/ variables
  - Values/ measures
- … that is labelled
  - Predefined classes
  - Predefined values

- Are we able to map the observations to their label?
  - 2 legs
  - Have feathers
  - Fly

Supervised Learning
What is it?

- Given a set of observations
  - Features/ variables
  - Values/ measures
- … that is labelled
  - Predefined classes
  - Predefined values

- Are we able to map the observations to their label?
  - 2 legs
  - Have feathers
  - Fly

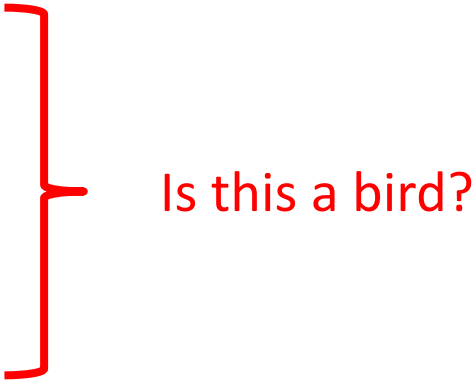
} Is this a bird?

Supervised Learning
What is it?

MONASH University

- Given a set of observations
  - Features/ variables
  - Values/ measures
- … that is labelled
  - Predefined classes
  - Predefined values

- Are we able to map the observations to their label?
  - 2 legs
  - Have feathers
  - Fly
  - Breathes
  - Sleep



Supervised Learning
What is it?

MONASH University

- Given a set of observations
  - Features/ variables
  - Values/ measures

} We want useful features only

- … that is labelled
  - Predefined classes
  - Predefined values

- Are we able to map the observations to their label?
  - 2 legs
  - Have feathers
  - Fly
  - ~~Breathes~~
  - ~~Sleep~~

} Is this a bird?

Supervised Learning
What is it?

MONASH University

- Given a set of observations
  - Features/ variables
  - Values/ measures
- … that is labelled
  - Predefined classes
  - Predefined values

- Are we able to map the observations to their values?
  - 700 square feet
  - 2 bed room
  - 2 bath room
  - LRT in 100 meter

Supervised Learning
What is it?

MONASH University

- Given a set of observations
  - Features/ variables
  - Values/ measures
- … that is labelled
  - Predefined classes
  - Predefined values

- Are we able to map the observations to their values?
  - 700 square feet
  - 2 bed room
  - 2 bath room
  - LRT in 100 meter

RM 500,000.00?

Supervised Learning
What is it?

- Can you think of some applications?

Supervised Learning
What is it?

MONASH University

- Can you think of some applications?
  - Organizing email

Supervised Learning
What is it?

MONASH University

- Can you think of some applications?
  - Organizing email
    - What is important
    - What is spam

Supervised Learning
What is it?

MONASH University

- Can you think of some applications?
  - Organizing email
    - What is important
    - What is spam
  - Classifying document
    - News
    - Sports
    - Social media

Supervised Learning
What is it?

MONASH University

- Can you think of some applications?
  - Organizing email
    - What is important
    - What is spam
  - Classifying document
    - News
    - Sports
    - Social media
  - High risk transactions
    - Credit card
    - Loan application

Supervised Learning
What is it?

MONASH University

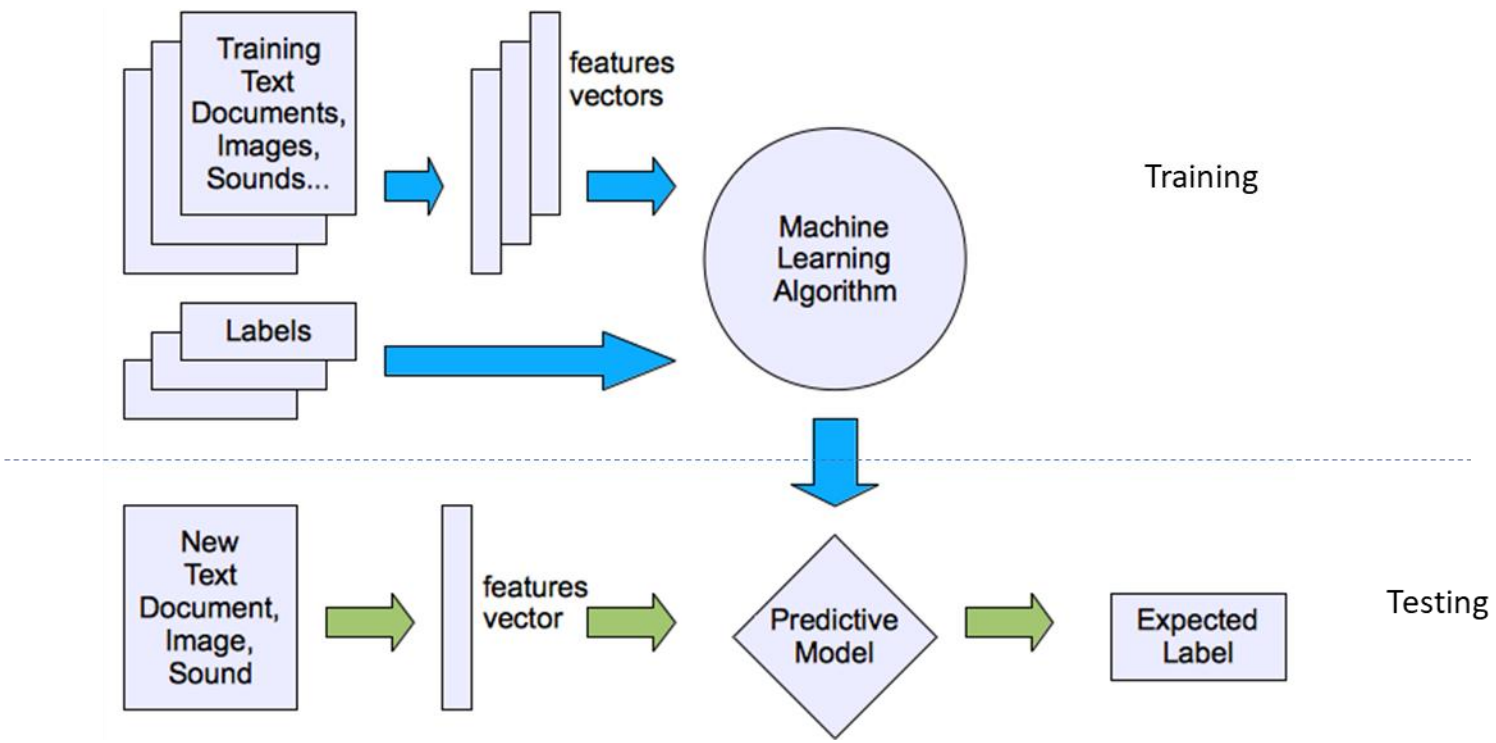
- Can you think of some applications?
  - Organizing email
    - What is important
    - What is spam
  - Classifying document
    - News
    - Sports
    - Social media
  - High risk transactions
    - Credit card
    - Loan application
  - High potential clients
    - Based on their spending
    - Based on their income etc.

MONASH University

Questions?

Supervised Learning

What is it?



MONASH University

Questions?

Supervised Learning
Algorithms

- Decision Tree
- Regression
- Naïve Bayes (NB)
- K-nearest Neighbour (kNN)
- Support Vector Machine (SVM)
- Artificial Neural Network (ANN)

MONASH University

- Decision Tree

MONASH University

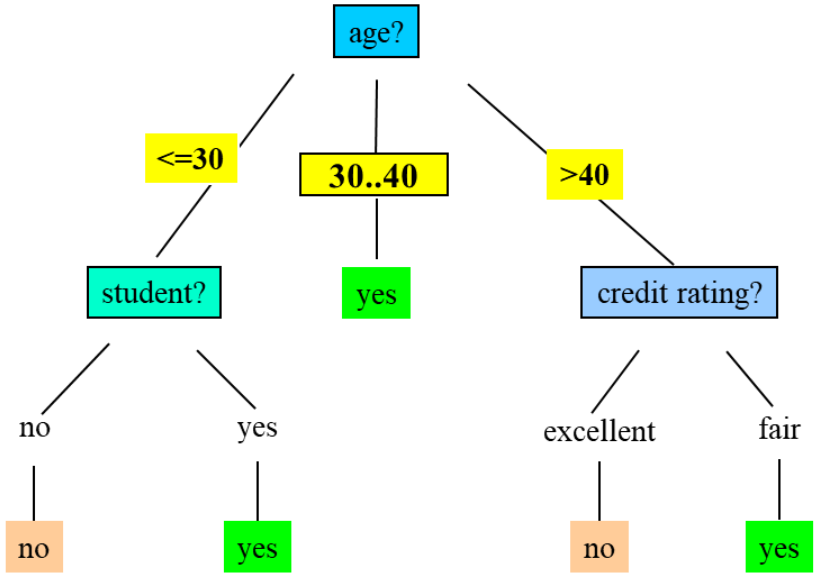
- Decision Tree

age	income	student	credit_rating	buys_computer
<=30	high	no	fair	no
<=30	high	no	excellent	no
31…40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31…40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31…40	medium	no	excellent	yes
31…40	high	yes	fair	yes
>40	medium	no	excellent	no

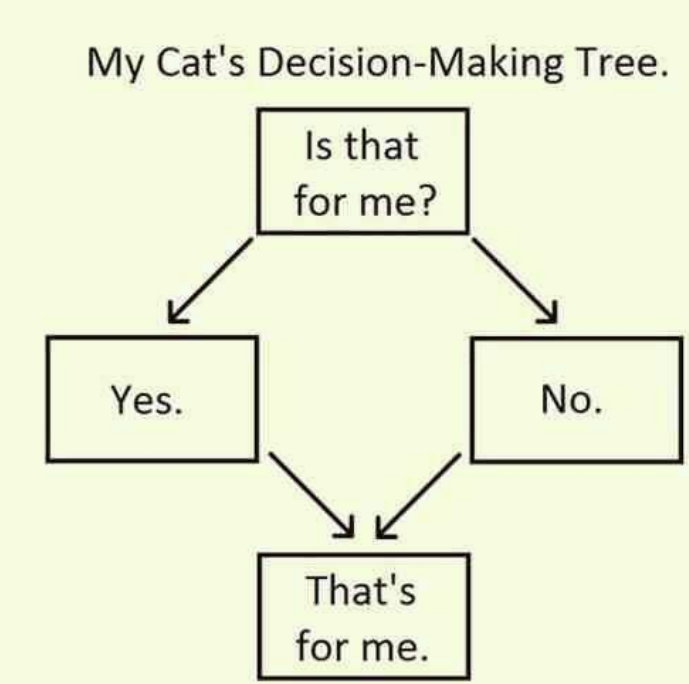
MONASH University

- Decision Tree

age	income	student	credit_rating	buys_computer
<=30	high	no	fair	no
<=30	high	no	excellent	no
31…40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31…40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31…40	medium	no	excellent	yes
31…40	high	yes	fair	yes
>40	medium	no	excellent	no



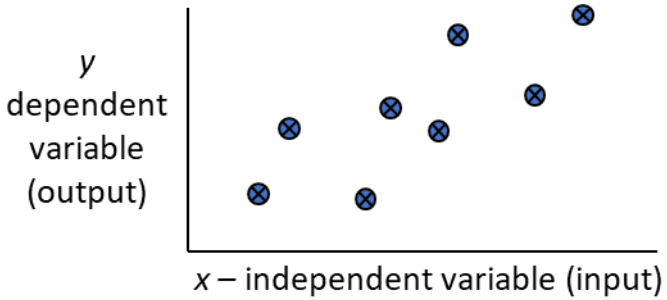
- Decision Tree



Supervised Learning
Algorithms

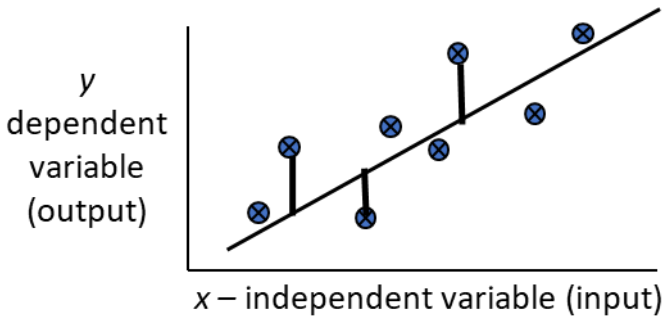
MONASH University

- Regression



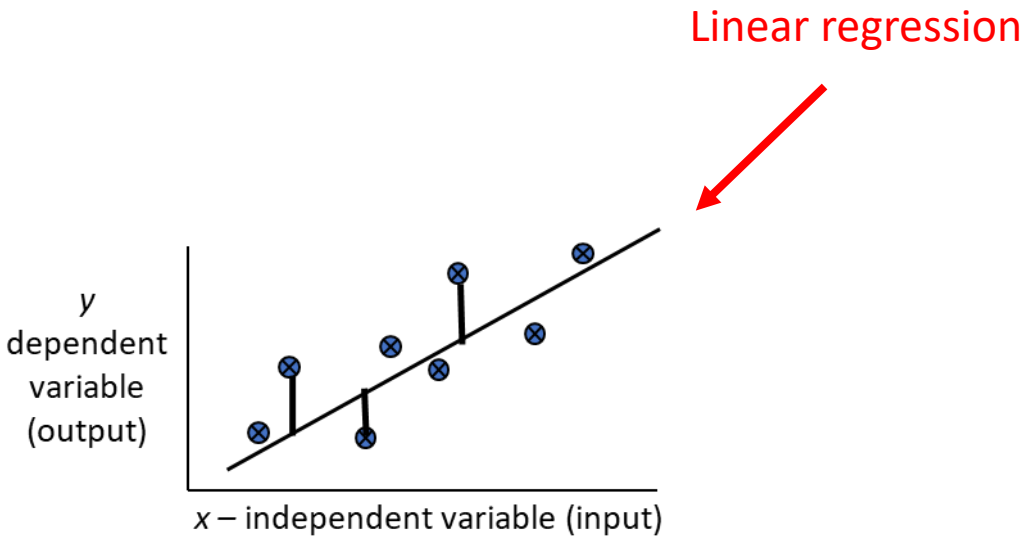
MONASH University

- Regression

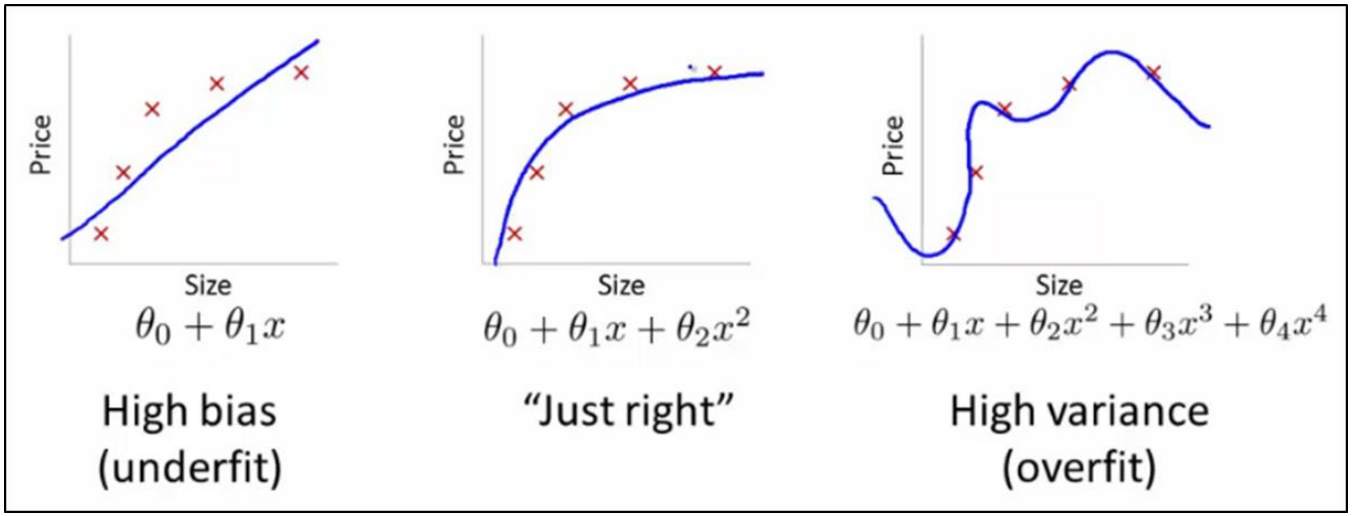


MONASH University

- Regression



- Regression


Price
Size
$\theta_0 + \theta_1 x$
High bias
(underfit)
Price
Size
$\theta_0 + \theta_1 x + \theta_2 x^2$
"Just right"
Price
Size
$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$
High variance
(overfit)

MONASH University

- Naïve Bayes (NB)

MONASH University

- Naïve Bayes (NB)
  - Probability-based using the Bayesian Theorem

$$P(C_j|A_1 \cap A_2 \cap A_3 \dots \cap A_n) = \frac{P(C_j) \cdot P(A_1 \cap A_2 \cap A_3 \dots \cap A_n|C_j)}{P(A_1 \cap A_2 \cap A_3 \dots \cap A_n)}$$

MONASH University

- Naïve Bayes (NB)
  - Probability-based using the Bayesian Theorem
  - Assuming variable independence

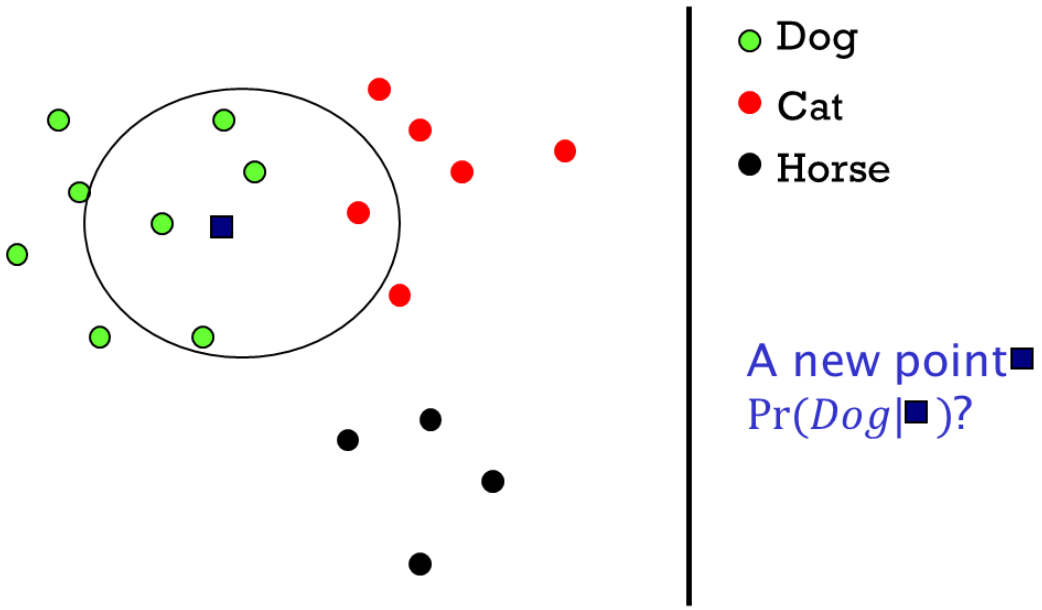
$$P(C_j | A_1 \cap A_2 \cap A_3 \dots \cap A_n) = \frac{P(C_j) \cdot P(A_1 \cap A_2 \cap A_3 \dots \cap A_n | C_j)}{P(A_1 \cap A_2 \cap A_3 \dots \cap A_n)}$$

MONASH University

- K-nearest Neighbour (kNN)

MONASH University

- K-nearest Neighbour (kNN)
  - When k=5, we look for the 5 closest points



Supervised Learning
Algorithms

- Support Vector Machine (SVM)
- Artificial Neural Network (ANN)

- We will skip these 2…

Supervised Learning
Algorithms

MONASH University

- Support Vector Machine (SVM)
- Artificial Neural Network (ANN)

- We will skip these 2…
- And even more complex ones
  - Ensemble methods for example

MONASH University

Questions?

Supervised Learning
Algorithms

MONASH University

- Let us try out a few simple examples
  - Decision tree
  - Regression

MONASH University

Questions?

MONASH University

- In the real world, it is more complex on how we do things… for example sentiment analysis

MONASH University

- Given a text, we pre-process it:

MONASH University

- Given a text, we pre-process it:
  - Convert case
  - Tokenization
  - Stop words removal
  - Stemming and Lemmatization
  - N-gram

MONASH University

- Given a text, we pre-process it:
  - “My dog eat my HomeWorK”
  - Convert case
  - Tokenization
  - Stop words removal
  - Stemming and Lemmatization
  - N-gram

MONASH University

- Given a text, we pre-process it:
  - "My dogs eated a HomeWorK"
  - Convert case = my dogs eated a homework
  - Tokenization = {my, dogs, eated, a, homework}
  - Stop words removal = {my, dogs, eated, homework}
  - Stemming and Lemmatization = {my, dog, eat, homework}
  - N-gram

MONASH University

- Given a text, we pre-process it:
  - "My dogs eated a HomeWorK"
  - Convert case = my dogs eated a homework
  - Tokenization = {my, dogs, eated, a, homework}
  - Stop words removal = {my, dogs, eated, homework}
  - Stemming and Lemmatization = {my, dog, eat, homework}
  - N-gram
    - 2-gram = {my dog} {dog eat} {eat homework}
    - 3-gram = {my dog eat} {dog eat homework}

MONASH University

- Given a text, we pre-process it:
  - "My dogs eated a HomeWorK"
  - Convert case = my dogs eated a homework
  - Tokenization = {my, dogs, eated, a, homework}
  - Stop words removal = {my, dogs, eated, homework}
  - Stemming and Lemmatization = {my, dog, eat, homework}
  - N-gram
    - 2-gram = {my dog} {dog eat} {eat homework}
    - 3-gram = {my dog eat} {dog eat homework}
- These are used as features…

MONASH University

- Given a text, we pre-process it:
  - "My dogs eated a HomeWorK"
  - Convert case = my dogs eated a homework
  - Tokenization = {my, dogs, eated, a, homework}
  - Stop words removal = {my, dogs, eated, homework}
  - Stemming and Lemmatization = {my, dog, eat, homework}
  - N-gram
    - 2-gram = {my dog} {dog eat} {eat homework}
    - 3-gram = {my dog eat} {dog eat homework}
- These are used as features…
  - Can we weighted such as TF-iDF

MONASH University

- Then we do the same with supervised learning

Supervised Learning
Sentiment Analysis

MONASH University

- Then we do the same with supervised learning
  - Positive sentiment sentences
  - Neutral sentences
  - Negative sentiment sentences

Supervised Learning
Sentiment Analysis

MONASH University

- Then we do the same with supervised learning
  - Positive sentiment sentences
  - Neutral sentences
  - Negative sentiment sentences
- Build a model which we can use!

MONASH University

- Let us try it out now!

MONASH University

- Let us try it out now!
- Of course real world is more complex
  - We bring in concepts as well

MONASH University

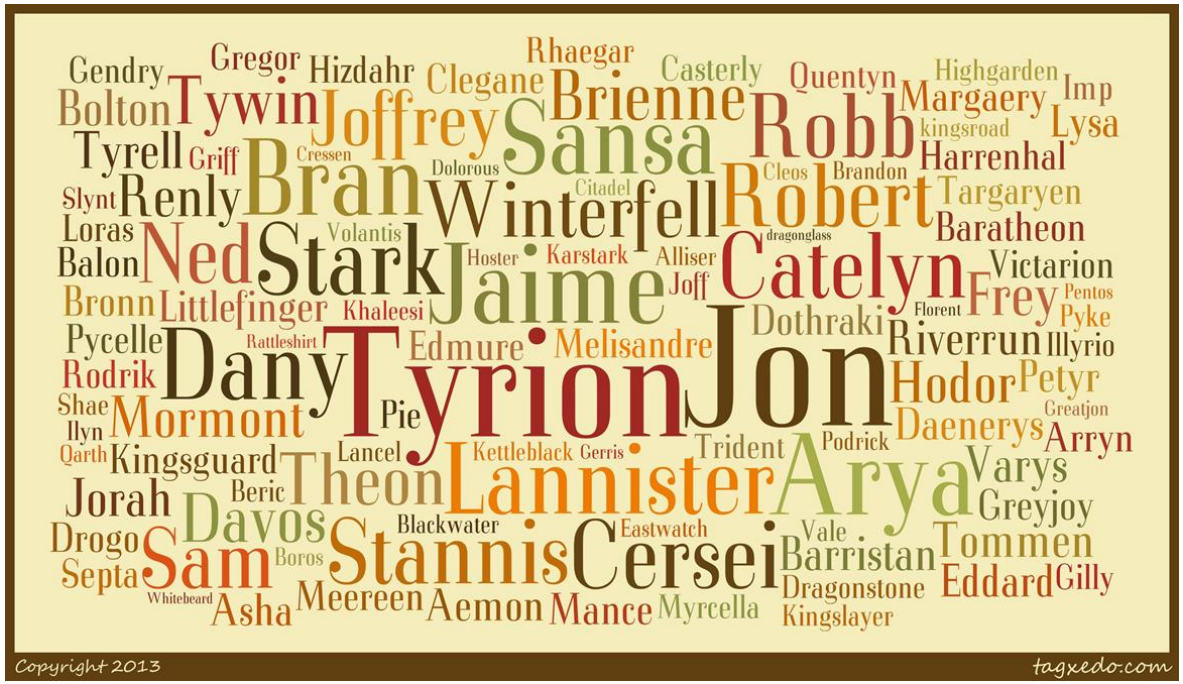
- Let us try it out now!

- Of course real world is more complex
  - We bring in concepts as well
  - For a hotel review
    - Food?
    - Location?
    - Room?
    - Price?

Supervised Learning

Sentiment Analysis

- Another usage is to understand important concepts in text such as word cloud



MONASH University

Questions?

MONASH University

Have a break

MONASH University

Questions?

Unsupervised Learning
What is it?

MONASH University

- Very similar with supervised learning…

Unsupervised Learning
What is it?

MONASH University

- Very similar with supervised learning…
- Except we are not giving
  - Predefined labels
  - Predefined values

Unsupervised Learning
What is it?

MONASH University

- Very similar with supervised learning…
- Except we are not giving
  - Predefined labels
  - Predefined values
- Then why do we do unsupervised learning?

Unsupervised Learning
What is it?

MONASH University

- Very similar with supervised learning…
- Except we are not giving
  - Predefined labels
  - Predefined values
- Then why do we do unsupervised learning?
  - Find association or relation
    - Features
    - Observations
  - Dimension reduction

Unsupervised Learning
What is it?

MONASH University

- Very similar with supervised learning…
- Except we are not giving
  - Predefined labels
  - Predefined values
- Then why do we do unsupervised learning?
  - Find association or relation
    - Features
    - Observations
  - Dimension reduction

} Can you think of applications?

Unsupervised Learning

What is it?

- Given a collection of food, how would you group it?

Unsupervised Learning

What is it?

- Given a collection of food, how would you group it?

Unsupervised Learning

What is it?

- Given a collection of human, how would you group it?

Unsupervised Learning
What is it?

- Given a collection of human, how would you group it?

Unsupervised Learning
What is it?

MONASH University

- Do be careful, it isn’t always predefined…

Unsupervised Learning
What is it?

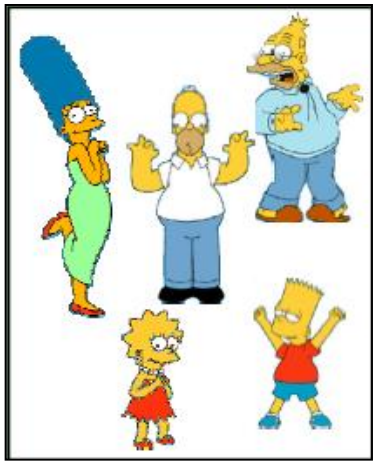
- Do be careful, it isn't always predefined…

Unsupervised Learning
What is it?

MONASH University

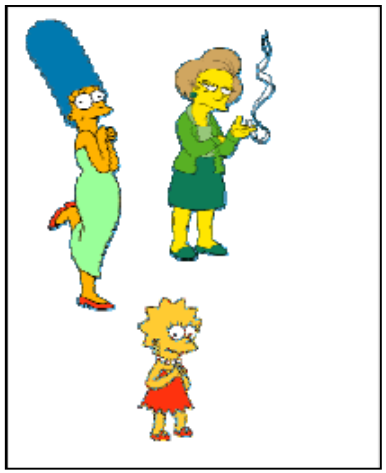
- Do be careful, it isn't always predefined…

Family



School

Female



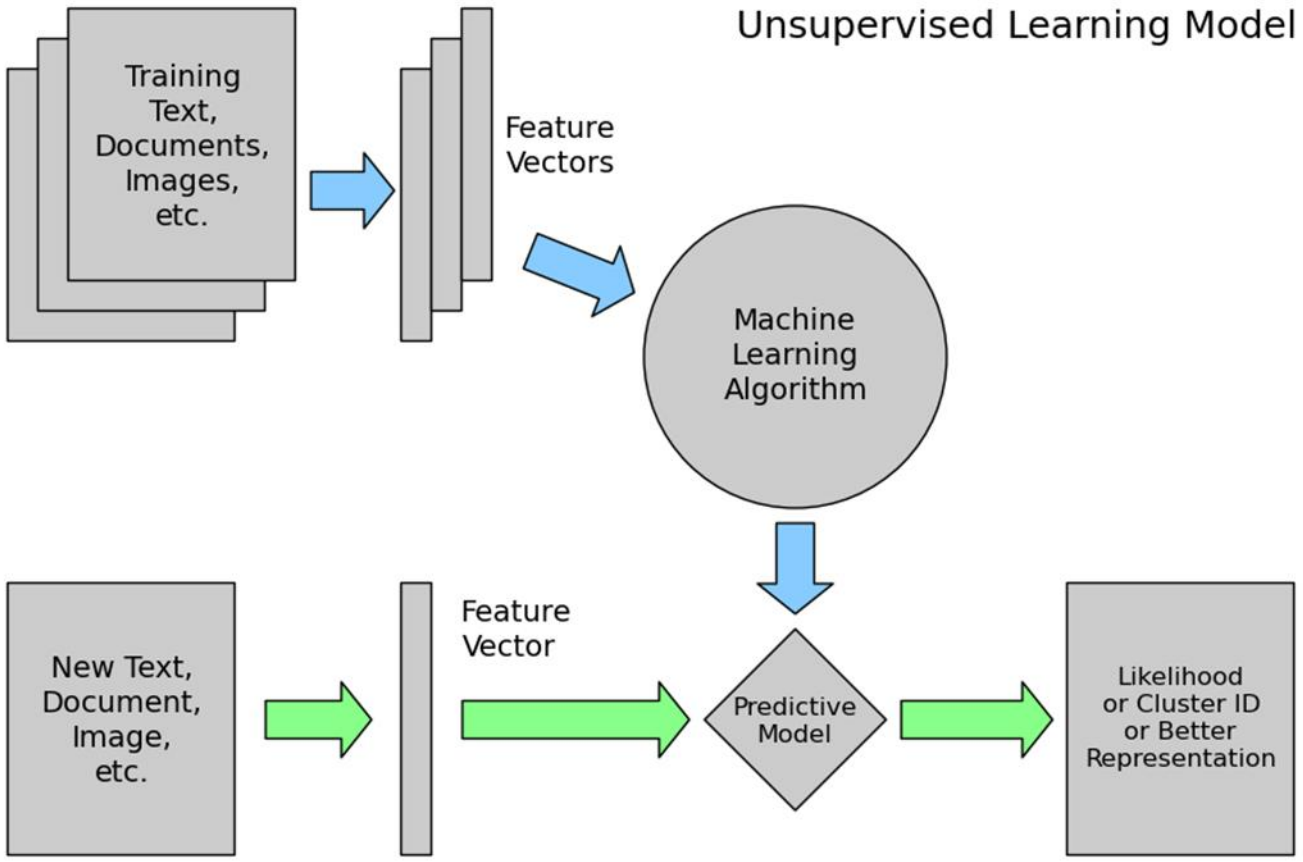
Male

MONASH University

Questions?

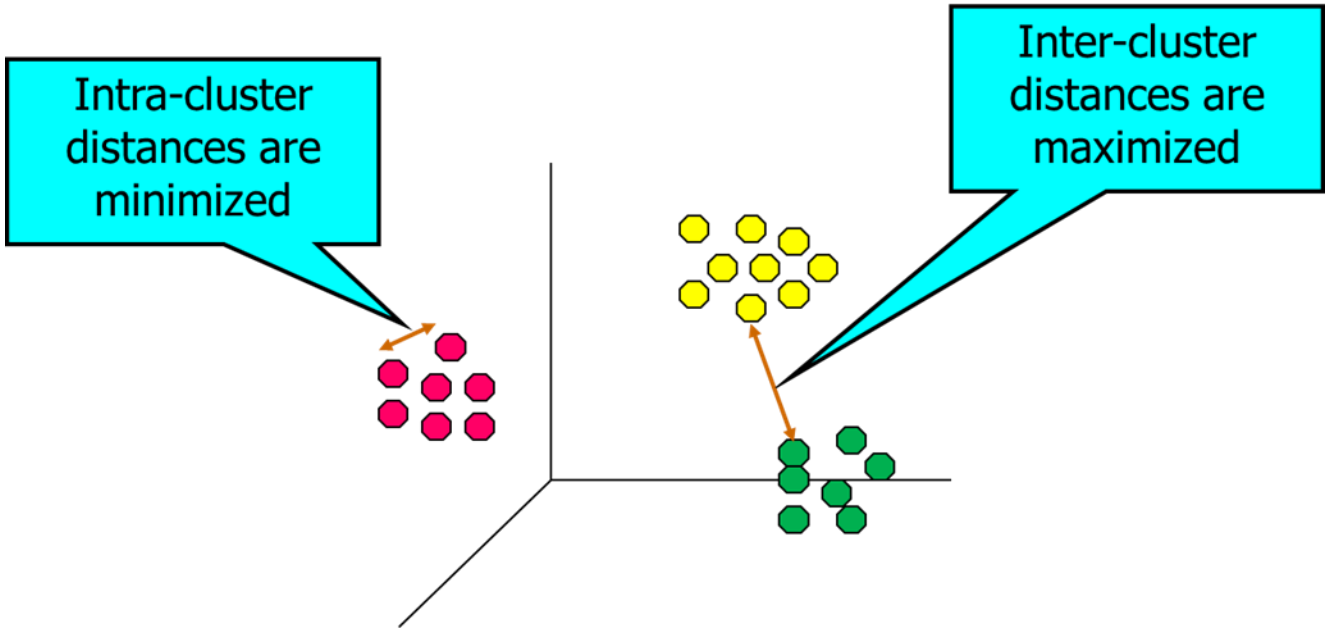
Unsupervised Learning

What is it?



Unsupervised Learning

What is it?



MONASH University

Questions?

Unsupervised Learning
Algorithms

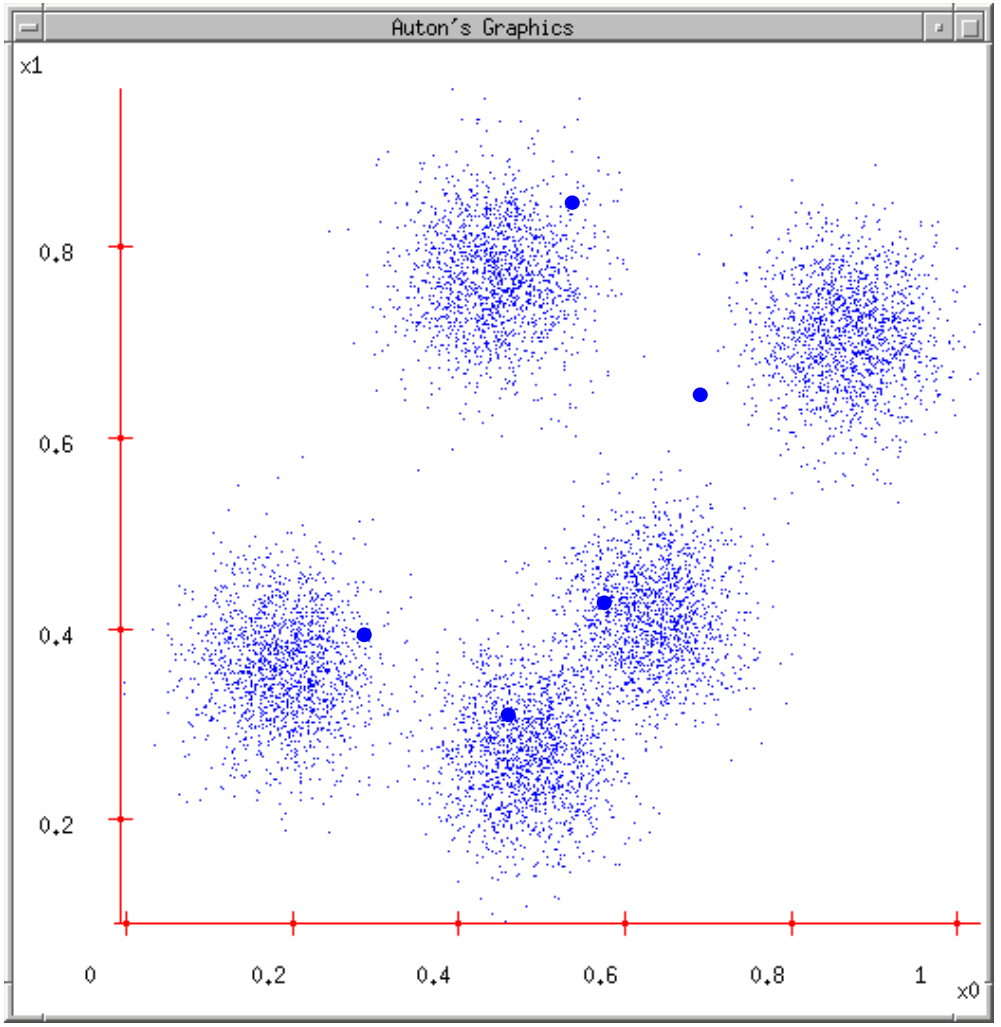
- K-means
- Gaussian Mixture Model (GMM)
- Mean Shift
- Hierarchical

Unsupervised Learning
Algorithms

- K-means

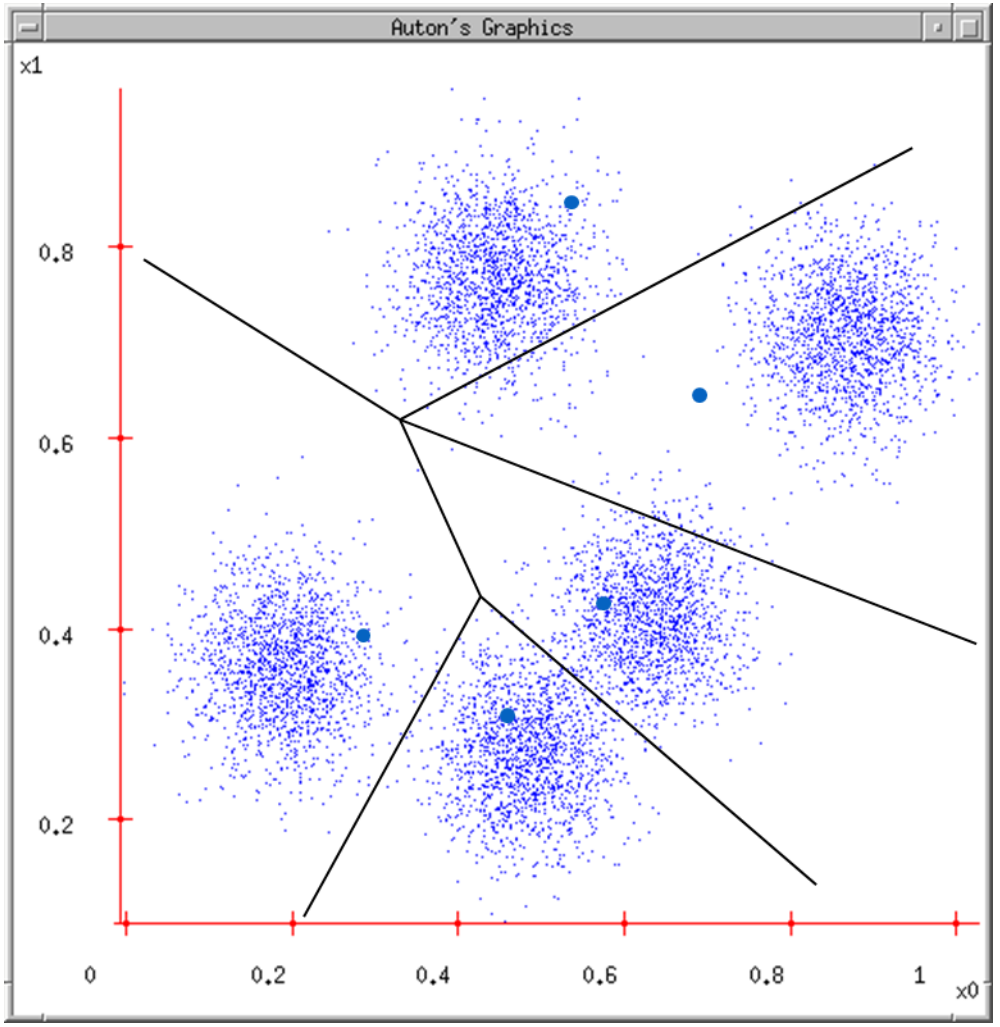
Unsupervised Learning
Algorithms

- K-means
  - We want 5 cluster



Unsupervised Learning
Algorithms

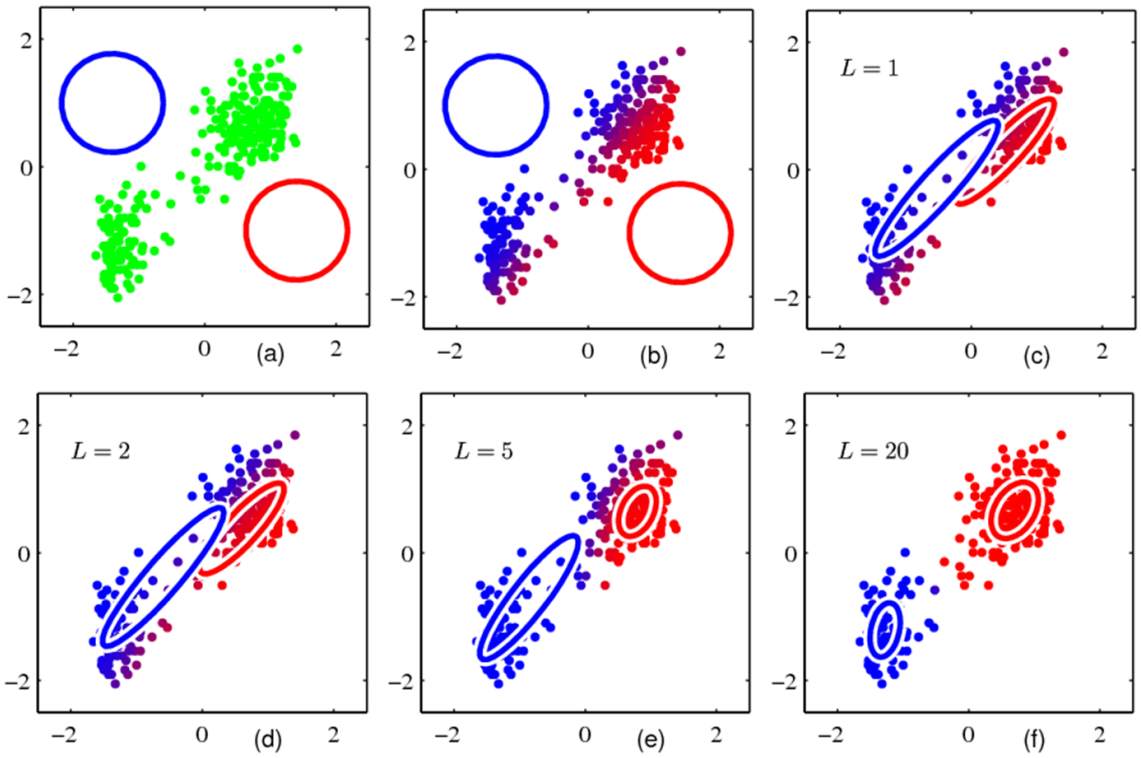
- K-means
  - We want 5 cluster
  - The process is more complex with calculations required



MONASH University

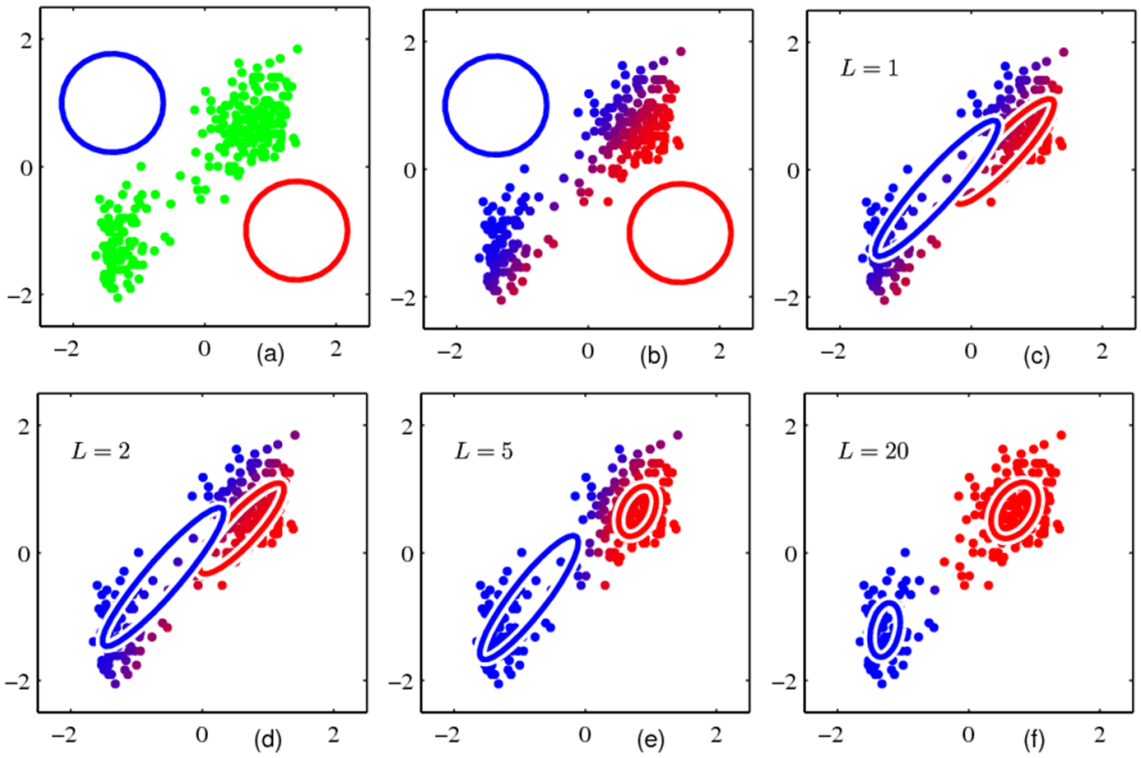
- Gaussian Mixture Model (GMM)

- Gaussian Mixture Model (GMM)
  - When it is more complex



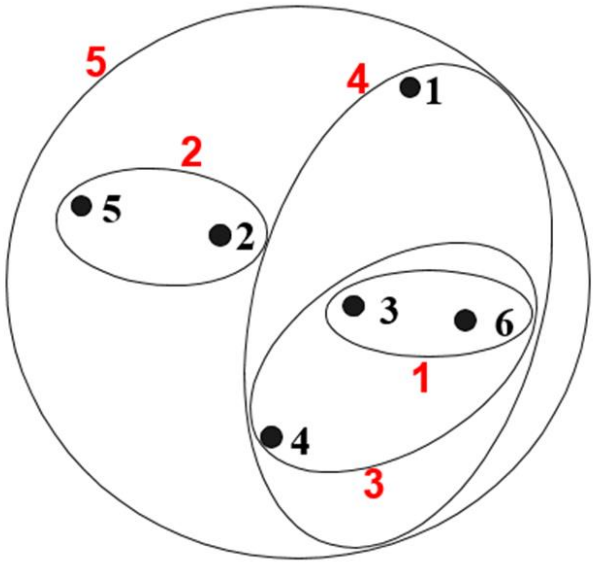
60

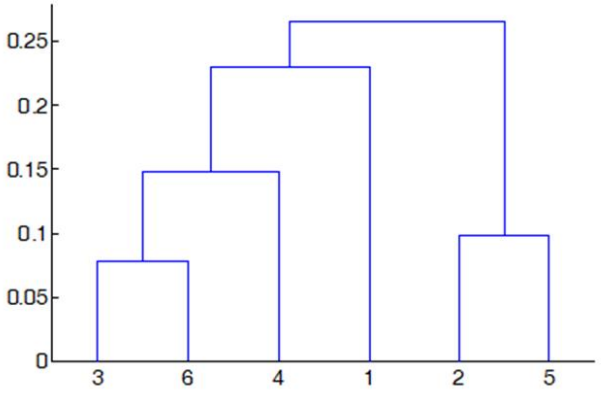
- Gaussian Mixture Model (GMM)
  - When it is more complex



60

- Hierarchical





MONASH University

- Mean Shift
  - We will skip this

MONASH University

- Mean Shift
  - We will skip this

- And there are more clustering algorithms out there, all grouping data in different way

MONASH University

- Mean Shift
  - We will skip this
- And there are more clustering algorithms out there, all grouping data in different way
  - For example the Latent Dirichlet Allocation (LDA) for topics

MONASH University

Questions?

- Our last hands-on

MONASH University

- Our last hands-on

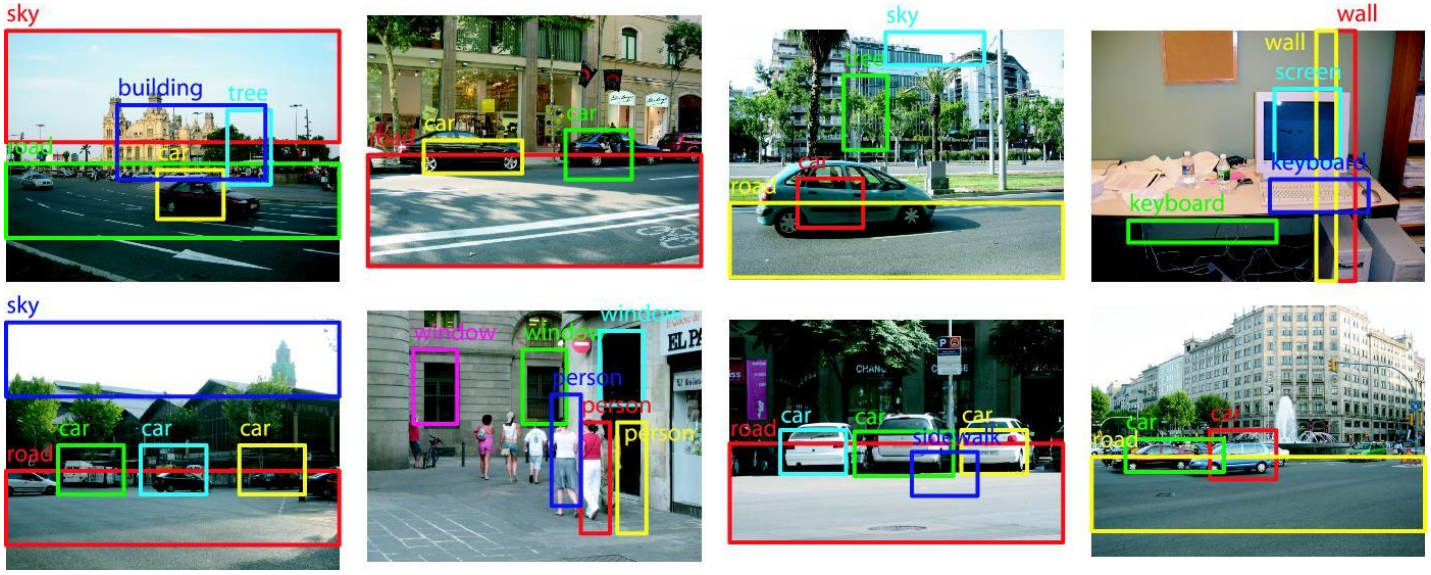
- Basically we are teaching the computer how to see things and recognize things

Computer Vision

Hands on…

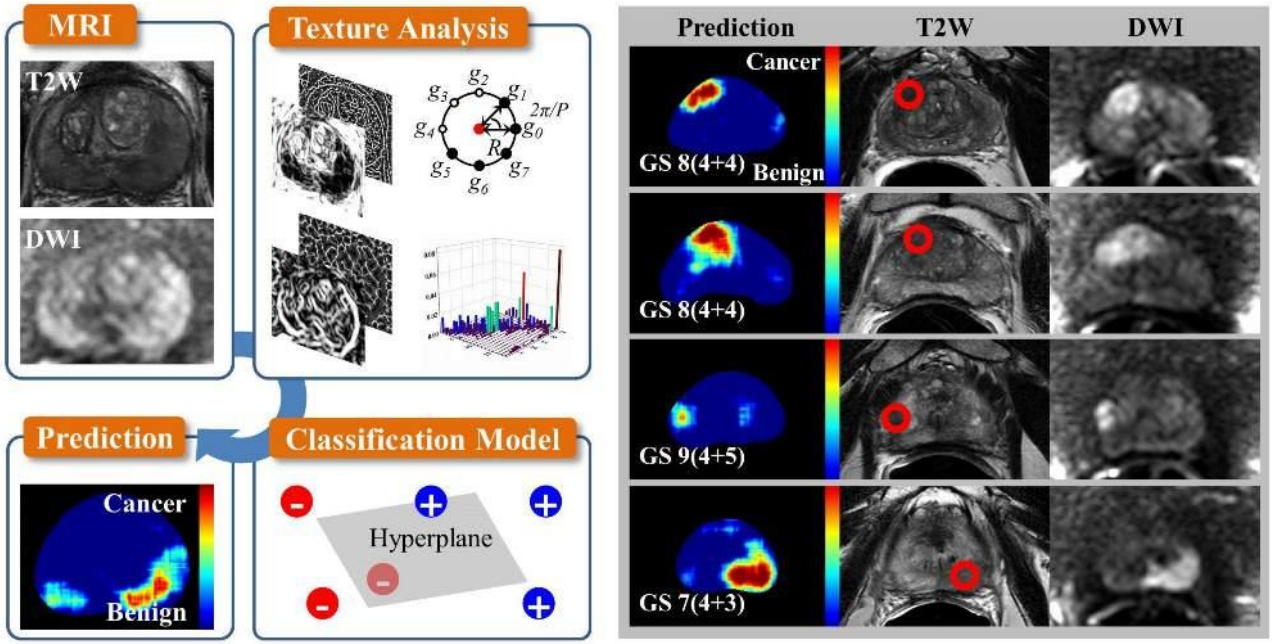
- Our last hands-on

- Basically we are teaching the computer how to see things and recognize things



- Our last hands-on

- Basically we are teaching the computer how to see things and recognize things



MONASH University

- Let us try some simple ones
  - Face
  - Eye
  - Smile

MONASH University

- Let us try some simple ones
  - Face
  - Eye
  - Smile

- But first, we need to learn how to extract features for the machine to recognize shapes
  - Edge detection as a simple example

MONASH University

Questions?

- It is useful

- It is useful
- Still untap potential

MONASH University

- It is useful
- Still untap potential
  - Good training?
  - Good features?
  - More features?

MONASH University

- It is useful
- Still untap potential
  - Good training?
  - Good features?
  - More features?
- Still a challenge

MONASH University

- It is useful
- Still untap potential
  - Good training?
  - Good features?
  - More features?
- Still a challenge
  - Overfitting
  - Underfitting

MONASH University

- It is useful
- Still untap potential
  - Good training?
  - Good features?
  - More features?
- Still a challenge
  - Overfitting
    - Imbalanced datasets
  - Underfitting

MONASH University

Questions?

MONASH University

Thank You