Machine Learning & Applications

IT3011

Stefan Winkler

NUS National University of Singapore

Credits: Pedro Domingos, Min-Yen Kan

1

Where Does Knowledge Come From?

Evolution

Culture

Experience

Computers

NUS National University of Singapore

2

"Most of the knowledge in the world in the future is going to be extracted by machines and will reside in machines."

Yann LeCun, Director of AI Research, Facebook

3

Knowledge Discovery by Computers

- Fill in gaps in existing knowledge
- Emulate the brain
- Simulate evolution
- Systematically reduce uncertainty
- Notice similarities between old and new

4

5 Tribes of ML

5 Tribes of Machine Learning

Tribe	Origins	Algorithm
Symbolists	Logic, Philosophy	Inverse deduction
Connectionists	Neuroscience	Backpropagation
Evolutionaries	Evolutionary biology	Genetic programming
Bayesians	Statistics	Probabilistic inference
Analogizers	Psychology	Support vector machines

Symbolists

7

Inverse Deduction

Addition	Subtraction
1	1
+ 1	+ ?
= ?	= 2

8

Inverse Deduction (Induction)

Deduction	Induction
Socrates is human	Socrates is human
+ Humans are mortal	+ ?
= ?	= Socrates is mortal

9

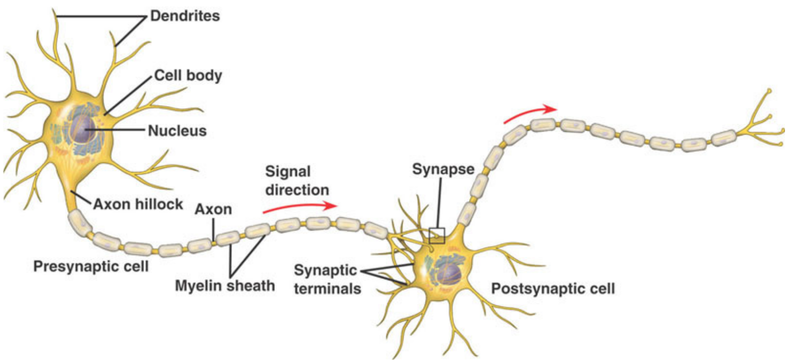
Robot Scientist (Drug Discovery)

10

Connectionists

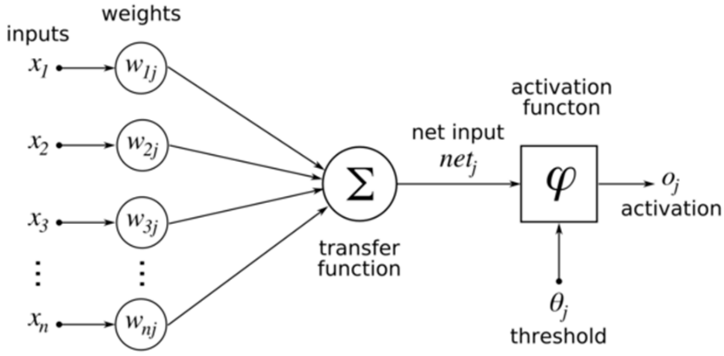
11

Neuron



12

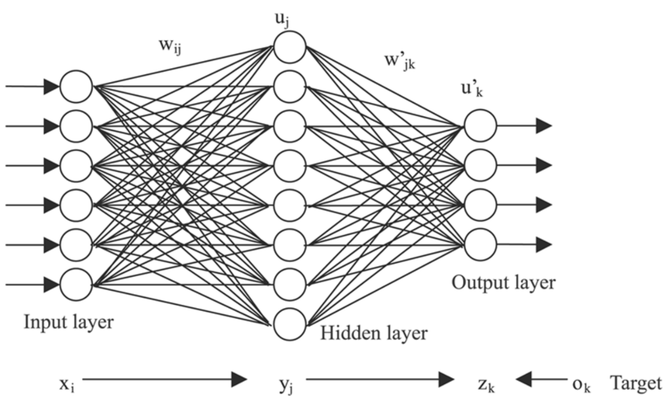
Artificial Neuron



IT3011: Machine Learning & Applications 13

13

Artificial Neural Network

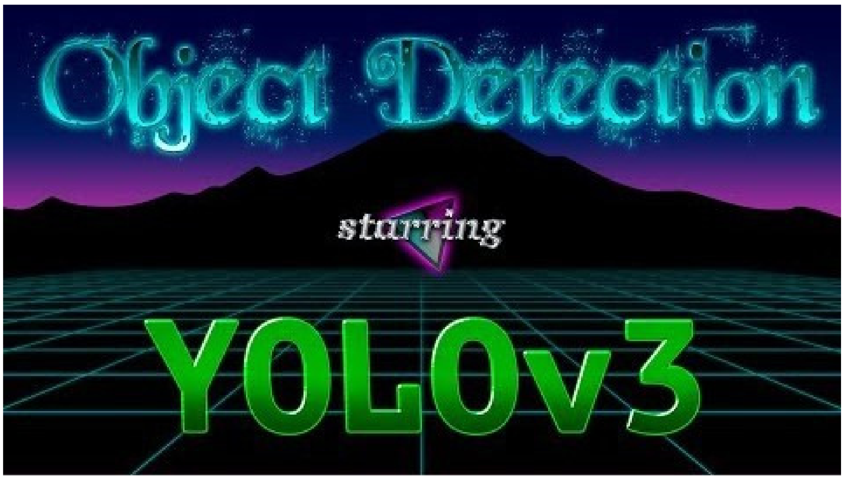


Determine weights via Backpropagation

IT3011: Machine Learning & Applications 14

14

Object Detection



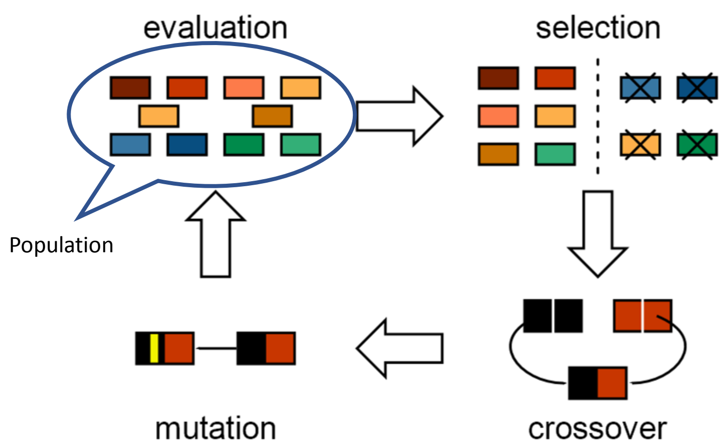
https://youtu.be/MPU2HistivI

15

Evolutionaries

16

Genetic Algorithms



17

Learning to Walk

https://vimeo.com/80822626

18

Bayesians

Probabilistic Inference

Likelihood: How probable is the data given that our hypothesis is true?

Prior: How probable was the hypothesis before observing the data?

$$P(y|X) = \frac{P(X|y)P(y)}{P(X)}$$

Posterior: How probable is our hypothesis given the observed evidence?

Marginal: How probable is the evidence under all possible hypotheses?

Spam Filtering

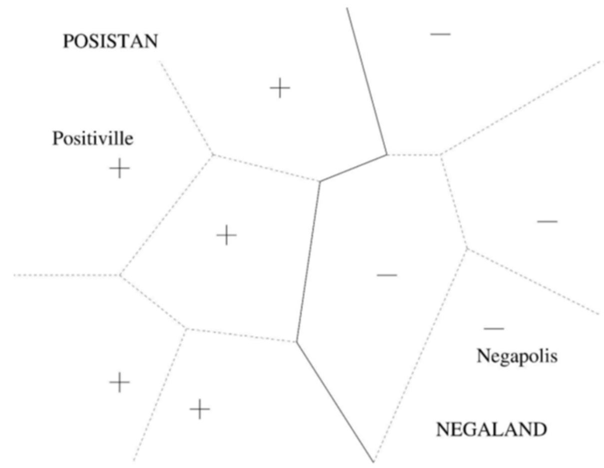
"Wow! I've got one from someone I know!"

21

Analogizers

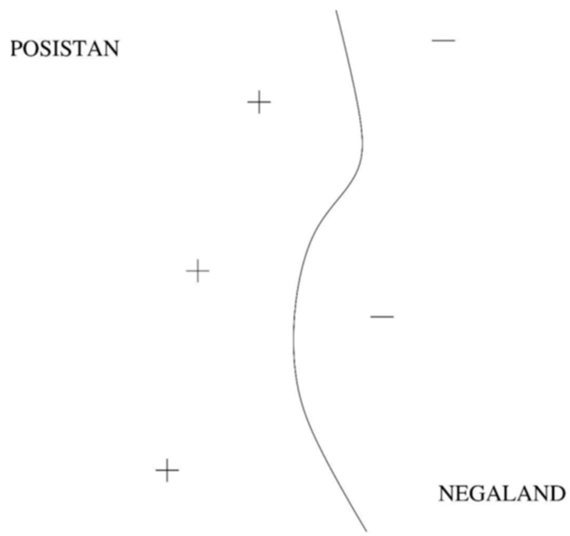
22

Nearest Neighbors



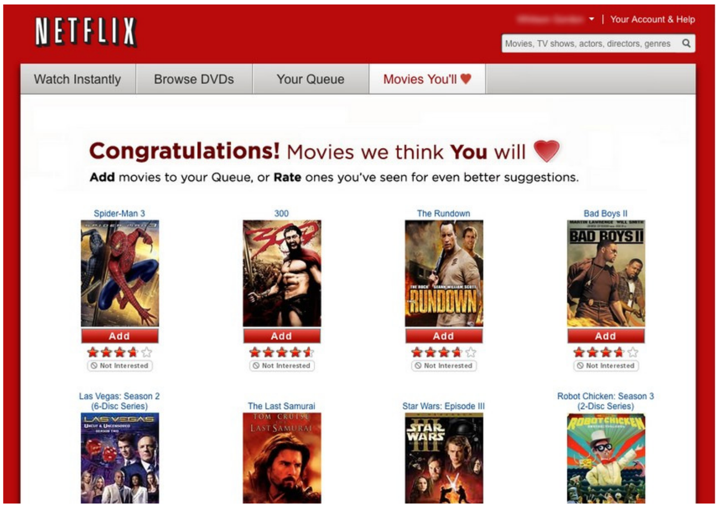
23

Support Vector Machines



24

Recommender Systems



IT3011: Machine Learning & Applications 25

25

5 Tribes of Machine Learning

Tribe	Problem	Algorithm
Symbolists	Knowledge Composition	Inverse deduction
Connectionists	Credit assignment	Backpropagation
Evolutionaries	Structure discovery	Genetic programming
Bayesians	Uncertainty	Probabilistic inference
Analogizers	Similarity	Support vector machines

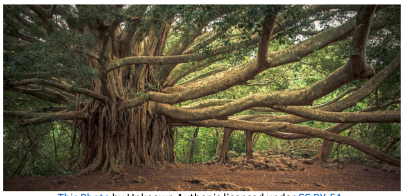
IT3011: Machine Learning & Applications 26

26

Back to Basics

27

What is a Tree?



- Toddlers can identify trees
- Defining a tree is not obvious

- 3 criteria that define ML well:
  - A pattern exists;
  - It's difficult to pin down formally (mathematically);
  - We have data for it.

28

Example: Credit Approval

- Banks need know to which loan applicants should be approved for their loan application.
  - If the bank approves an applicant that eventually defaults on their loan, this is a loss for the bank.
  - Conversely, if they reject an applicant that could actually pay back the loan with the compounded interest, they have missed an opportunity to make money.
- Input: Application
- Output: Decision

  - A pattern exists;
  - It's difficult to pin down formally (mathematically);
  - We have data for it.

Example: Credit Approval

- Inputs (observations, variables, features):

Variable / Feature	Value
Age	26
Gender	Female
Salary	$50k
Debt	$250k
Credit score	643

  - *n*-dimensional vector $\boldsymbol{x}$

- Outputs (labels, ground truth, targets):
  - Binary variable *y = {approve, reject}*
- <Input, output> pair = instance

Notation

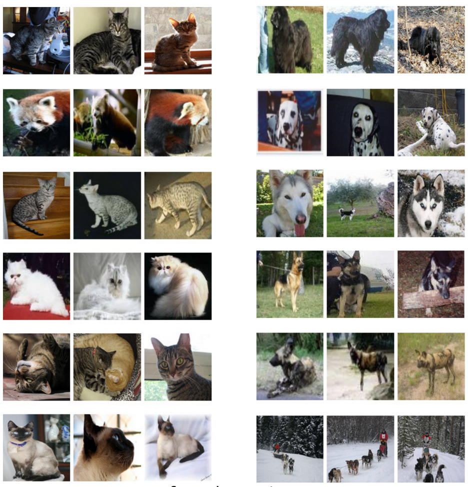
Notation	Meaning
$\mathbf{x}$	an input example (vector)
$\mathbf{X}$	a stacked set of inputs of form $\mathbf{x}$
$\mathbf{y}$	set of (aligned) outputs
n	number of features
m	number of instances
$\mathbf{x}_i$	i^{th} feature of input $\mathbf{x}$
$\mathbf{x}^{(j)}$	j^{th} input or instance

- Data matrix **X**: Collection of all input vectors (rows of **X**)

IT3011: Machine Learning & Applications 31

31

Supervised ML



Source: image-net.org

cat dog

- Function *f(**x**)* ⟶ *y*
- *Learn best possible approximation*, $h_\theta(\mathbf{x})$
- Classification vs. regression
  - A pattern exists;
  - It's difficult to pin down formally (mathematically);
  - We have data for it.
- *What else?*

IT3011: Machine Learning & Applications 32

32

Python example: sklearn (scikit-learn)

- Using colab
- https://www.comp.nus.edu.sg/~winkler/IT3011/01b.colab

33

Unsupervised ML

Source: image-net.org

- *How/what can you learn without labels?*
- Clustering
  - Grouping the data such that data points in the same group are **similar** to other data points in the same group and dissimilar to data points in other groups

34

Where is Machine Learning Useful?

- Automatic pattern discovery in large datasets
- Domains where humans lack knowledge to define effective features/algorithms
- Dynamically adapting to changing conditions

Issues in Machine Learning

- What algorithms exist for learning from training examples?
- In what settings will particular algorithms converge?
- Which algorithms performs best?
- How much training data is necessary/sufficient?
- What bounds on accuracy/confidence can be determined?
- How to best choose the next data point for training?
- When/how can prior knowledge guide the process?
- What specific function should the system learn?