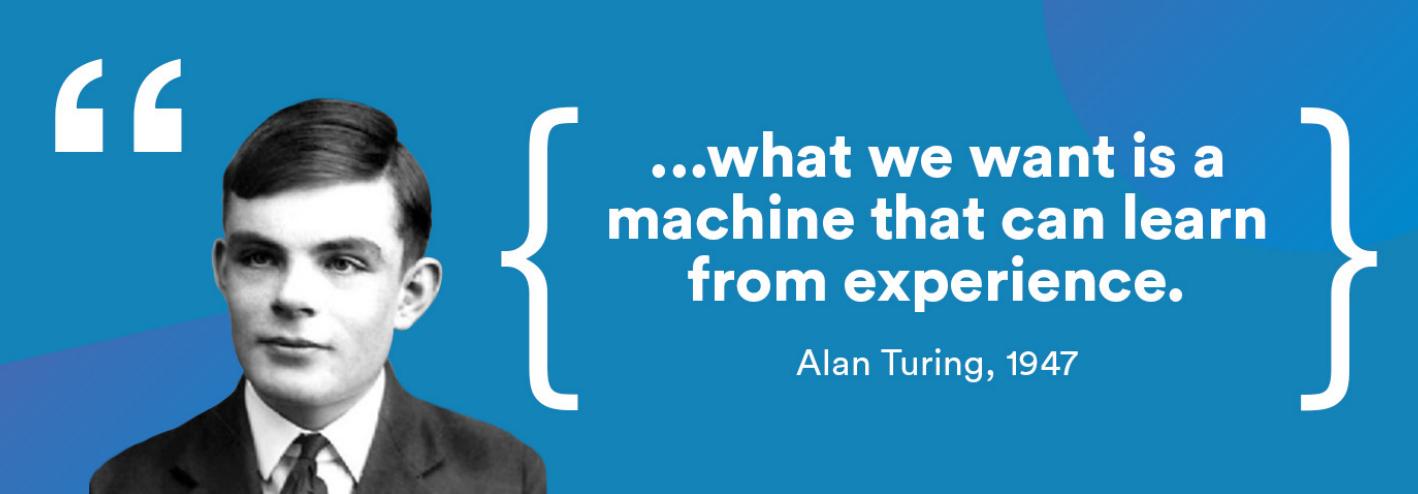
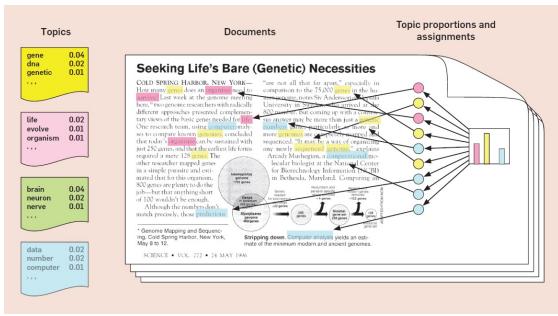
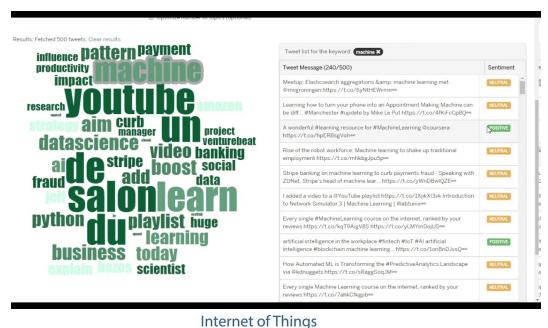
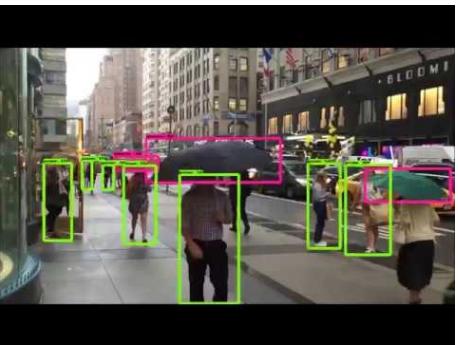


Introduction to Machine Learning

Dr. Srijith P K



- “A breakthrough in machine learning would be worth ten Microsofts” (Bill Gates, Chairman, Microsoft)
- “Machine learning is the next Internet” (Tony Tether, Director, DARPA)
- “AI is one of the most important things humanity is working on. It is more profound than, I don't know, electricity or fire.” - Sundar Pichai (Google CEO)



Recommender Systems



Self-driving cars



UBER

UBER eats



Machine Learning at Uber

- Personalized Application
- Estimated Time of Arrival

Auto friend tagging suggestion
Facebook suggests if you want to tag the person in the pic

Ads Recommendation
Recommends ads based on your search history
Machine Learning is used in generating recommendation
35 % of Amazon's revenue is generated by its recommendation system



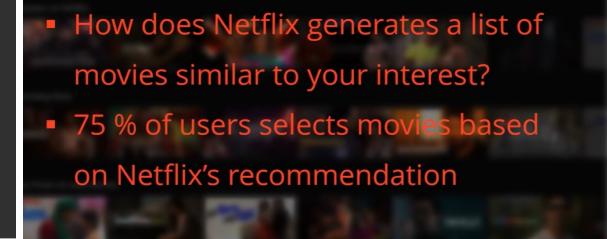
SCANNING

- Driverless Cars!
- Tesla's AI is driven by Nvidia's H/W focusing mainly on unsupervised learning
- Crowdsource data from all of its vehicles and its drivers - internal & external sensors



NETFLIX

Recommender System



- How does Netflix generates a list of movies similar to your interest?
- 75 % of users selects movies based on Netflix's recommendation

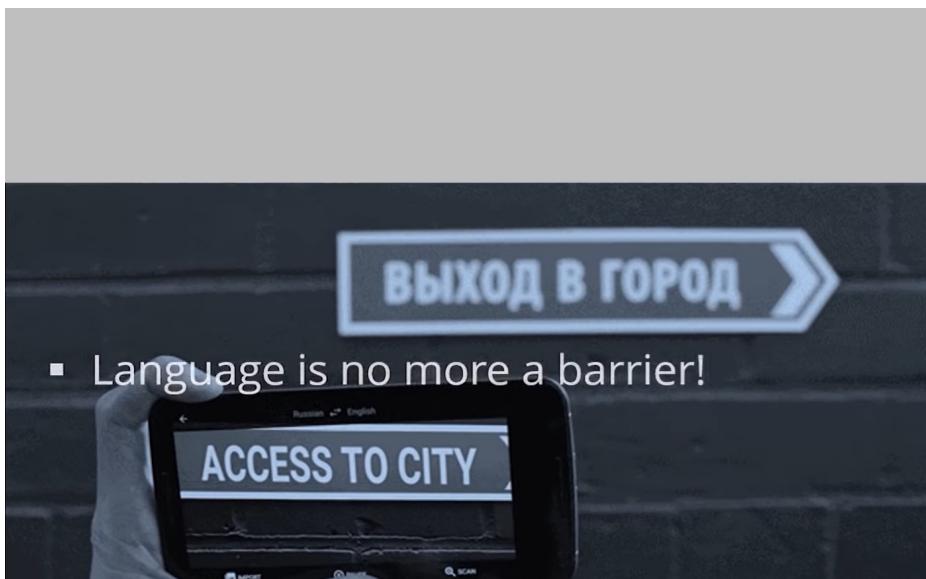
Machine Learning at Apple

- Smartphone with facial recognition
- Core of the face detection - machine learning algorithms

iPhone



- Language is no more a barrier!



Google Translate
Break through language barriers



- Google: 3.5 billion search queries every day.²
- Facebook: 350 million photos are uploaded to Facebook each day. Facebook generates 4 petabytes of data every day.
- Every day, 306.4 billion emails are sent, and 5 million Tweets are made in Twitter:
- Astronomy: Satellite data is in hundreds of PB.



twitter

<https://techjury.net/blog/how-much-data-is-created-every-day/#gref>



shutterstock.com • 593204357

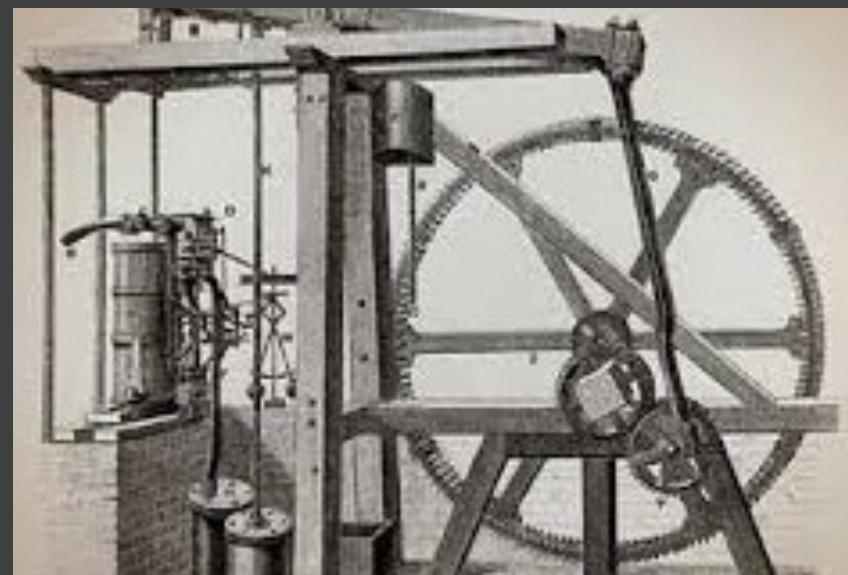
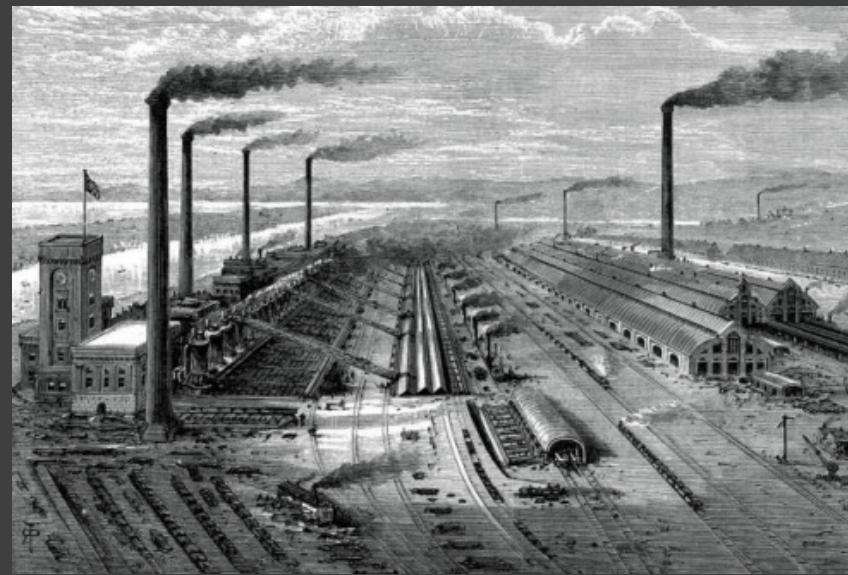
Big data Era

- **1.7MB of data** is created every second by every person during 2020.
- In the last two years alone, the astonishing **90%** of the world's data has been created.
- **2.5 quintillion bytes** of data are produced by humans every day.
- **463 exabytes** of data will be generated each day by humans as of 2025.
- **95 million** photos and videos are shared every day on Instagram.
- By the end of 2020, **44 zettabytes** will make up the entire digital universe.
- Every day, **306.4 billion emails** are sent, and **5 million Tweets** are made.

Why machine learning : Data every where !

"We are drowning in information and starving for knowledge." – John Naisbitt.

- This deluge of data calls for automated methods of data analysis, which is what machine learning provides.
- Defined as a “set of methods that can automatically detect patterns in data, and then use the uncovered patterns to predict future data, or to perform other kinds of decision making under uncertainty” - Kevin Murphy (Machine Learning : A Probabilistic Perspective)

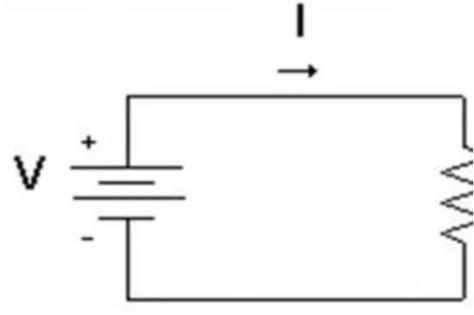




Whats Machine
Learning ?

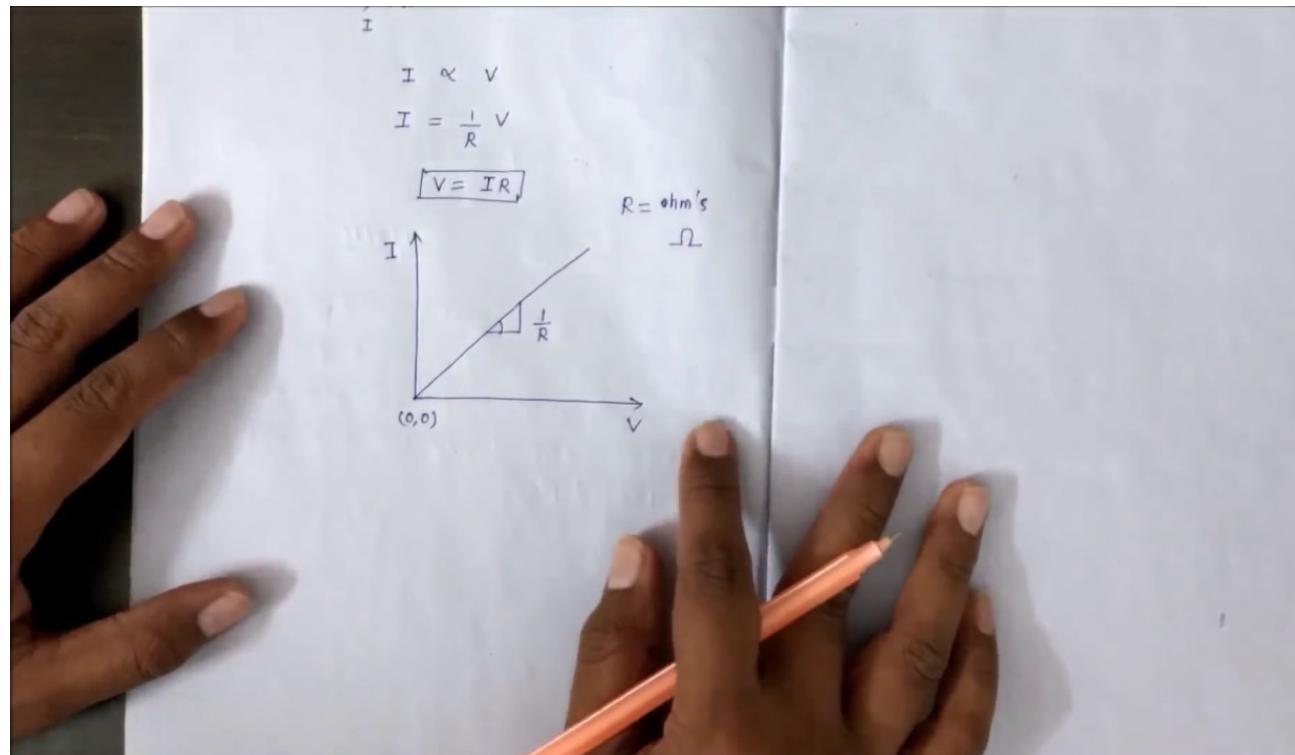
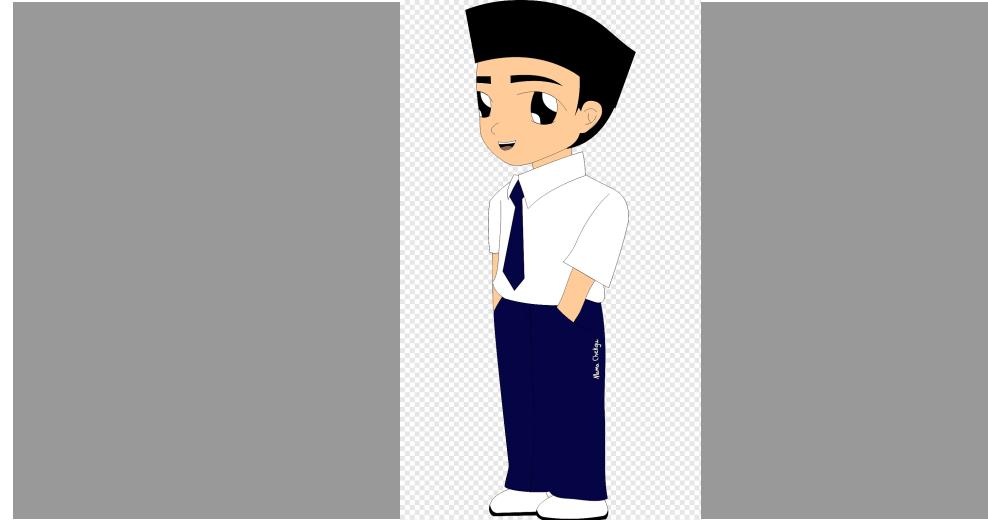


Georg Simon Ohm



Basic Electrical Circ

$$I = \frac{V}{R}$$



Estimate Resistance

- Connect the variable voltage supply to both the ends of the rheostat. Connect the ammeter in series of the rheostat. Connect the voltmeter in parallel of the rheostat. Start measuring the voltage and current as you move the rheostat moving hand from minimum position to the maximum position in the steps of constant increase in current.

Potential difference V (in volt)	0.5	1.0	1.5	2.0	2.5
Current I (in ampere)	0.2	0.4	0.6	0.8	1.0

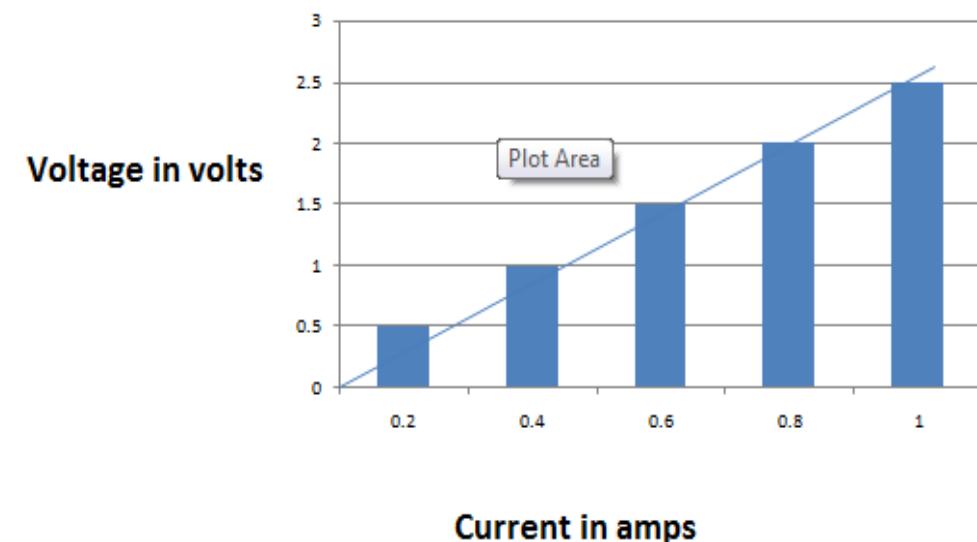
Estimating Resistance is machine learning !

- A currentvoltage characteristic or IV curve (currentvoltage curve) is a relationship, typically represented as a chart or graph, between the electric current through a circuit, device, or material, and the corresponding voltage, or potential difference across it. In the graph, the voltage is plotted along the y-axis and the current is plotted along the x-axis.

$$V = R I$$

resistance is 2.5 ohms!

Potential difference V (in volt)	0.5	1.0	1.5	2.0	2.5
Current I (in ampere)	0.2	0.4	0.6	0.8	1.0



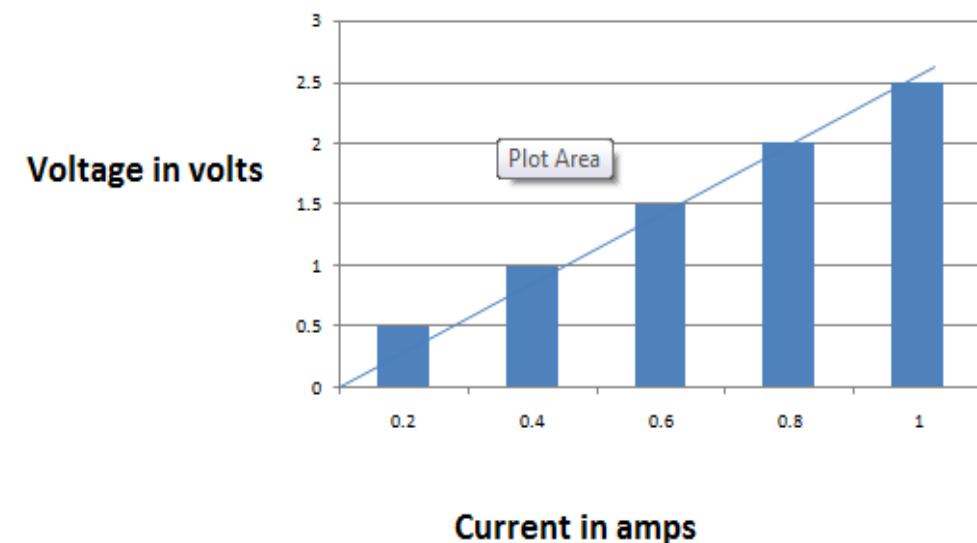
Estimating Resistance is machine learning !

- A currentvoltage characteristic or IV curve (currentvoltage curve) is a relationship, typically represented as a chart or graph, between the electric current through a circuit, device, or material, and the corresponding voltage, or potential difference across it. In the graph, the voltage is plotted along the y-axis and the current is plotted along the x-axis.

$$V = R I$$
$$y = m x$$

Output Input

Potential difference V (in volt)	0.5	1.0	1.5	2.0	2.5
Current I (in ampere)	0.2	0.4	0.6	0.8	1.0



What's machine learning ?

Human Learning at the age of 6 months.



Converged at the age of 12 months



Transfer skills at the age of 14 months



- Algorithms which let machines learn like humans from observations !
- To discover the fundamental principles of learning from data and use them to develop algorithms that can learn like living beings !
- Programming computers to optimize a performance criterion using example data or past experience(Ethem Alpaydin, Machine Learning, 2010)
- How do we create computer programs that improve with experience? (Tom Mitchel)

ML based applications

-Machine learning has become prominent approach to solve problems in AI domains like computer vision, language and speech processing

The screenshot shows the Google Translate interface on a web browser. The search bar at the top contains "google translate". Below it, a search result for "are you feeling down" is shown, with the English text "are you feeling down" on the left and its Hindi translation "क्या आप नीचे महसूस कर रहे हैं" on the right. Below the translation, the text "kya aap neeché mahasoos kar rahe hain" is displayed. The interface includes standard Google search navigation links like "All", "Books", "News", "Maps", "Images", "More", "Settings", and "Tools". A note at the bottom states: "Google Translate https://translate.google.com/ Google's free service instantly translates words, phrases, and web pages between English and over 100 other languages."

Jeopardy! (2011): Humans vs. IBM Watson



By Rosemarytoufee (Own work), via Wikimedia Commons

Natural Language Understanding and information extraction!

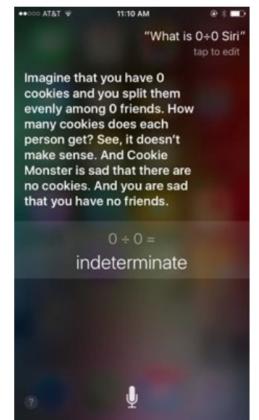
Face detection



Viola-Jones method.

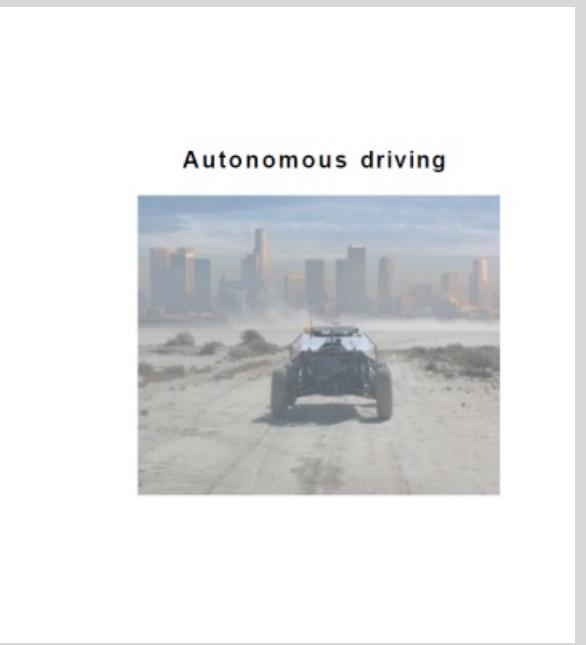
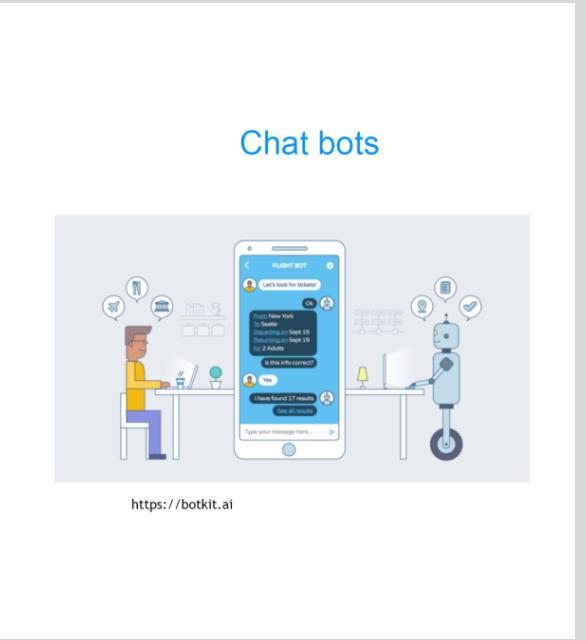
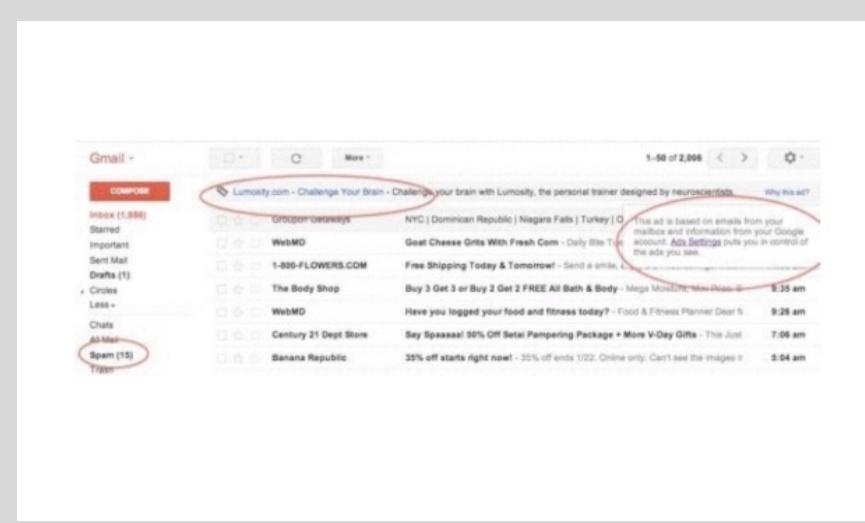
Speech recognition

- Virtual assistants: Siri (Apple), Echo (Amazon), Google Now, Cortana (Microsoft).
- “They” helps get things done: send an email, make an appointment, find a restaurant, tell you the weather and more.
- Leverage deep neural networks to handle **speech recognition** and **natural language understanding**.



ML based applications

- Machine learning has become prominent approach to solve problems in AI domains like computer vision, language and speech processing
- Early approaches to AI was based on **logic** but applications has to face a lot of uncertain situations and has to perform well on unseen situations.
- Machine learning focused on developing algorithms which could perform well on future unseen data (**generalization performance**) which differentiates it from statistics



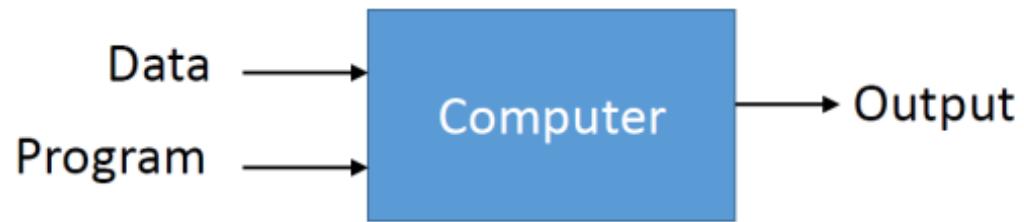
Machine learning is interdisciplinary

- Science (Astronomy, neuroscience, medical imaging, bio-informatics)
- Environment (energy, climate, weather, resources)
- Retail (Intelligent stock control, demographic store placement)
- Manufacturing (Intelligent control, automated monitoring, detection methods)
- Security (Intelligent smoke alarms, fraud detection)
- Marketing (promotions, ...)
- Management (Scheduling, timetabling)
- Finance (credit scoring, risk analysis...)
- Web data (information retrieval, information extraction, ...)



Machine learning

Traditional Programming



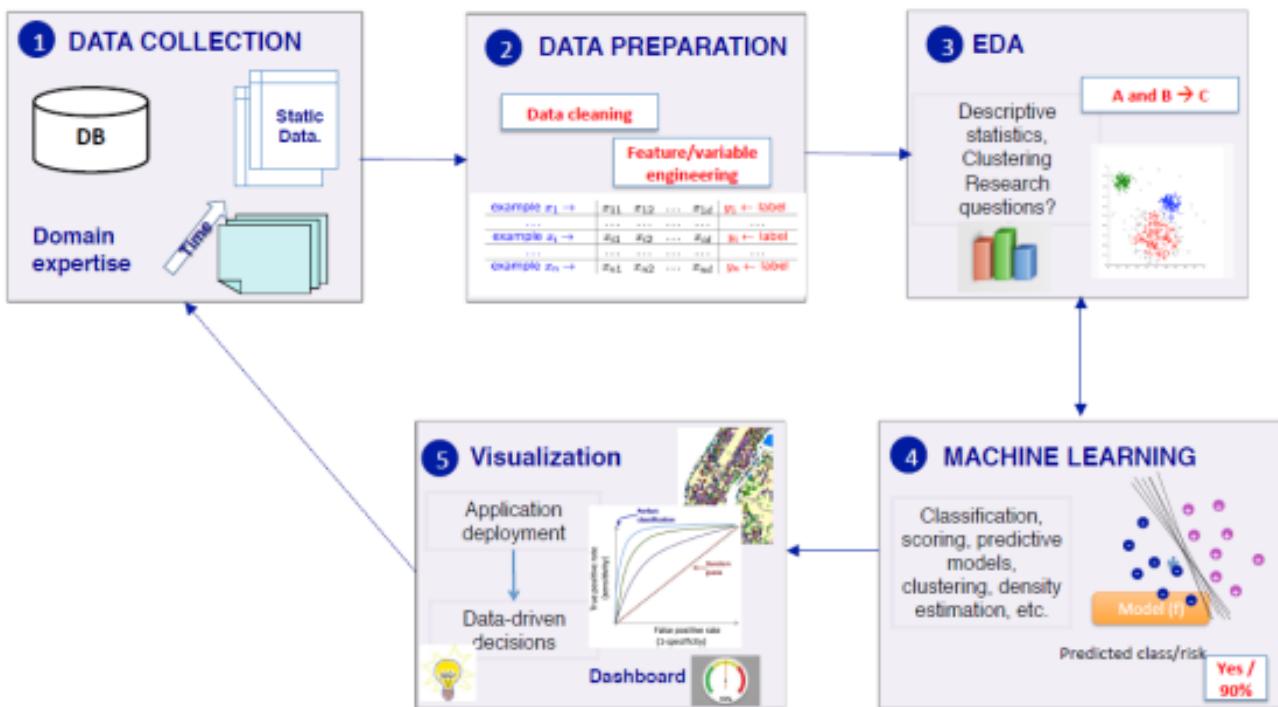
Machine Learning



Machine Learning, Data Mining, Knowledge Discovery,
Artificial Intelligence, Statistical Learning, Pattern Recognition,
Computational Learning



ML in practice



- Understanding domain, prior knowledge, and goals
- Data integration, selection, cleaning, pre-processing, etc.
- Learning models
- Interpreting results
- Consolidating and deploying discovered knowledge
- Loop

Overview of Machine learning

Supervised learning

- Predict an output y when given an input x
- For categorical y : classification.
- For real-valued y : regression.

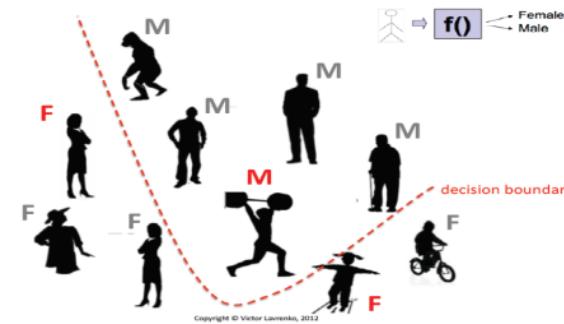
Unsupervised learning

- Create an internal representation of the input, e.g. clustering, dimensionality reduction
- This is important in machine learning as getting labels is often difficult and expensive

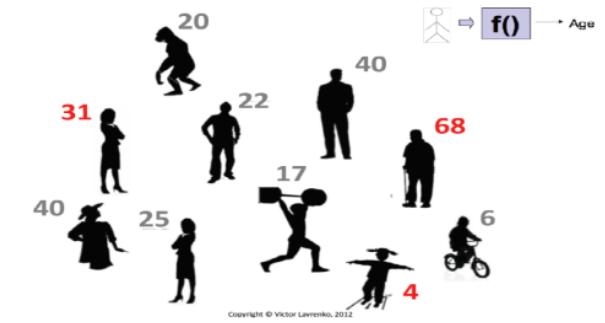
Other settings of ML

- Reinforcement learning (learning from “rewards”)
- Semi-supervised learning (combines supervised + unsupervised)
- Active learning, online learning, Transfer learning, multi-task learning, Structured prediction

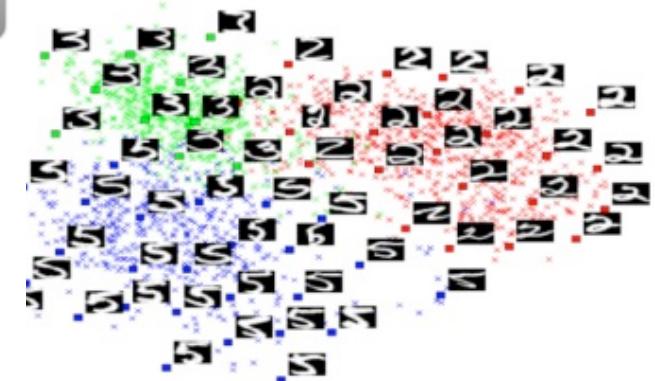
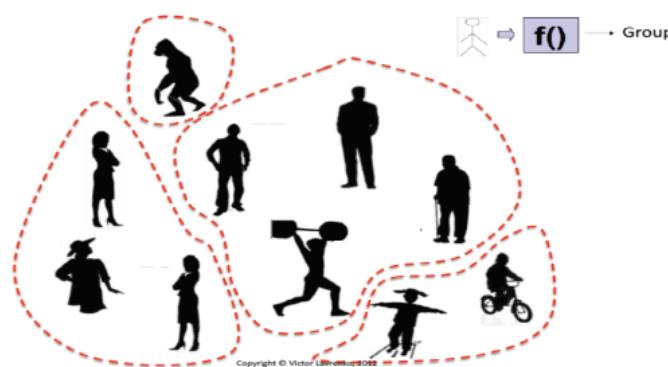
Classification (Supervised Learning)



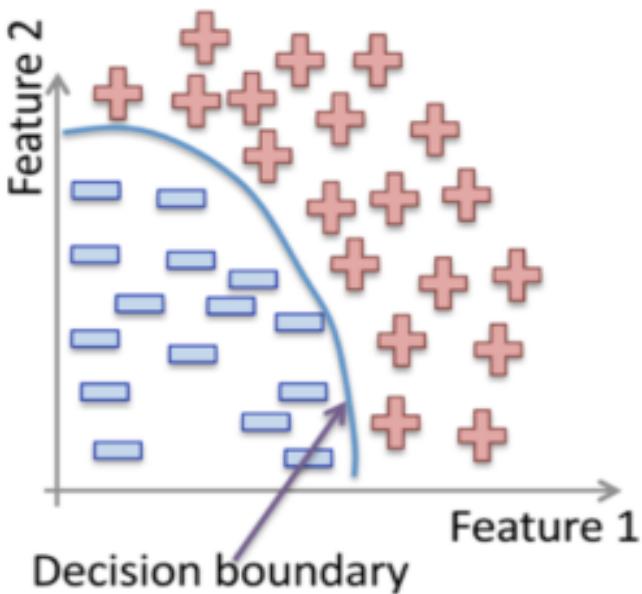
Regression (Supervised Learning)



Clustering (Unsupervised Learning)



Supervised learning (classification)



Input $\mathbf{x} \in \mathcal{R}^d$ and output y a label. Learn a function

$$f : \mathbf{x} \rightarrow y$$

Input

example $x_1 \rightarrow$ $x_{11} \quad x_{12} \quad \dots \quad x_{1d} \quad y_1 \leftarrow \text{label}$

...

example $x_i \rightarrow$ $x_{i1} \quad x_{i2} \quad \dots \quad x_{id} \quad y_i \leftarrow \text{label}$

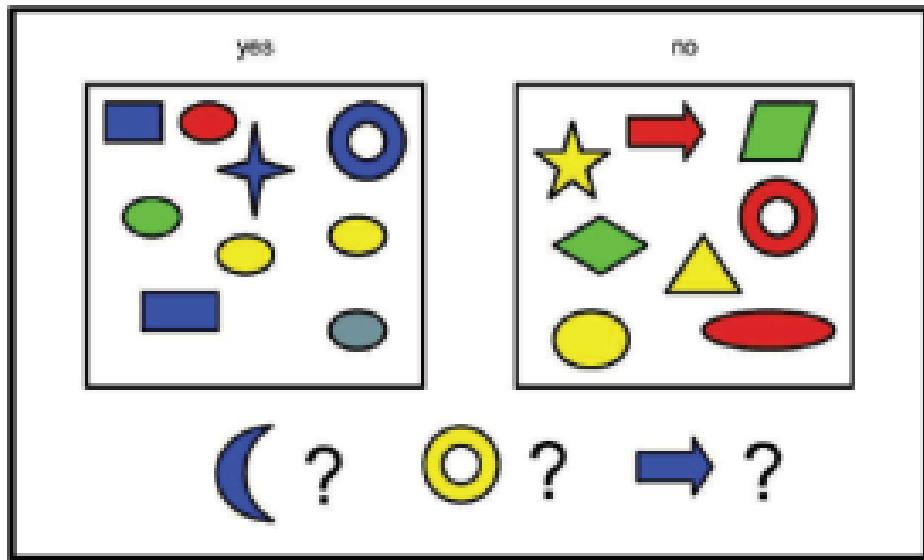
...

example $x_n \rightarrow$ $x_{n1} \quad x_{n2} \quad \dots \quad x_{nd} \quad y_n \leftarrow \text{label}$

Output

fruit	length	width	weight	label
fruit 1	165	38	172	Banana
fruit 2	218	39	230	Banana
fruit 3	76	80	145	Orange
fruit 4	145	35	150	Banana
fruit 5	90	88	160	Orange
...
fruit n

Supervised learning (Classification)

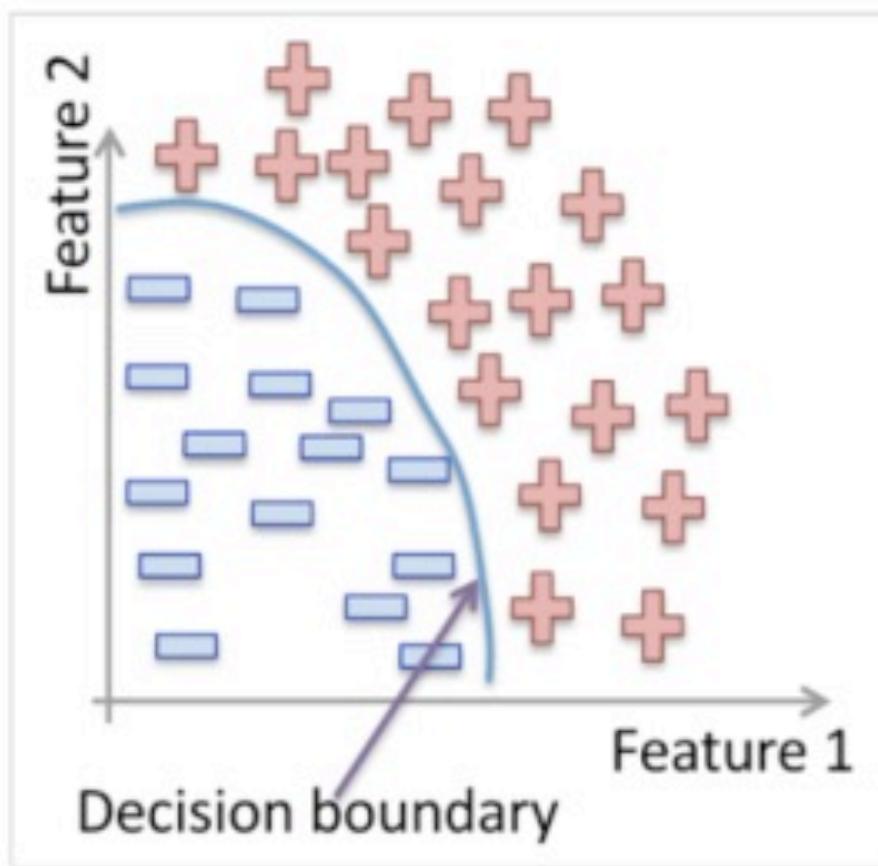


D features (attributes)

N cases

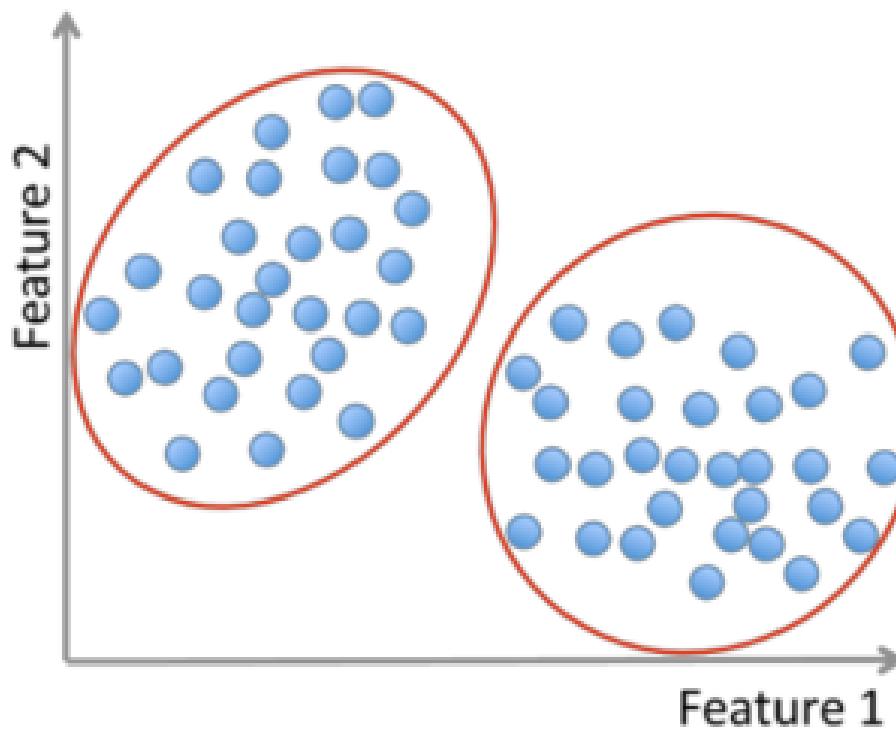
Color	Shape	Size (cm)	Label
Blue	Square	10	1
Red	Ellipse	2.4	1
Red	Ellipse	20.7	0

Supervised learning



Methods: Support Vector Machines, neural networks, decision trees, K-nearest neighbors, naive Bayes, etc.

Unsupervised learning

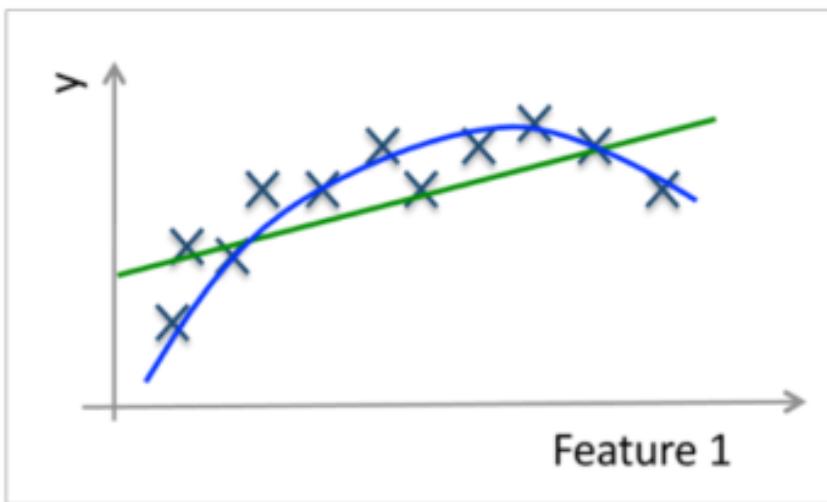


Methods: K-means, gaussian mixtures, hierarchical clustering, spectral clustering, etc.

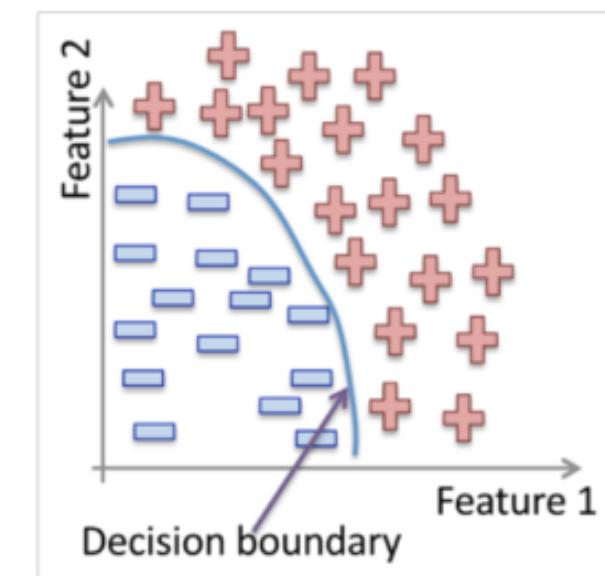


Machine learning algorithms

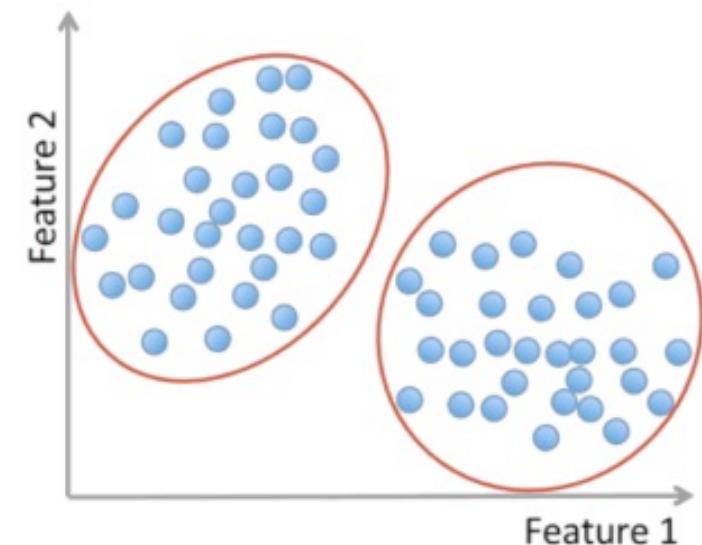
Regression:



Method : Linear regression, support vector regression, gaussian process regression



Methods: Support Vector Machines, neural networks, decision trees, K-nearest neighbors, naive Bayes, etc.



Methods: K-means, gaussian mixtures, hierarchical clustering, spectral clustering, etc.

Brief Syllabus : Foundations of ML

- Overview of machine learning
- Probability Models
- Supervised learning : Regression and Classification
 - Linear/Nonlinear regression
 - Logistic regression
 - K nearest neighbors
 - Naïve Bayes Classifier
 - Decision trees
 - Support vector machines
 - Neural Networks
 - Ensemble Methods
 - Model Selection and regularization
 - Bias and Variance
- Unsupervised Learning
 - Clustering : K-Means, Gaussian mixture models
 - Hierarchical clustering
 - Dimensionality Reduction : PCA, kernel PCA
- Probabilistic Graphical Models
- Reinforcement Learning
- Multi-task learning, online learning, transfer learning, active learning

Machine learning

Main references

- Mitchell, Tom. Machine Learning. New York, NY: McGraw-Hill, 1997
- Alpaydin, Ethem Introduction to Machine Learning. MIT Press, 2014.
- Bishop, Christopher M. Pattern Recognition and Machine Learning. Springer, 2006.
- Murphy, K. P. (2013). *Machine learning : a probabilistic perspective*. Cambridge, Mass. MIT Press
- Hastie, T., R. Tibshirani, and J. H. Friedman. The Elements of Statistical Learning: Data Mining, Inference and Prediction. New York, NY: Springer, 2001

ML Resources

- MOOCs
 - Coursera, EdX, Udacity
- Conferences/Journals
 - JMLR, Machine Learning, IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions on Pattern Analysis and Machine Intelligence, Annals of Statistics
 - ICML, NIPS, KDD, IJCAI, AAAI, ICDM

ML Datasets

- UCI Repository:
<http://www.ics.uci.edu/~mlearn/MLRepository.html>
- Statlib: <http://lib.stat.cmu.edu/>
- Kaggle
- Many more...

Foundations of ML

- Assume mathematical and programming basics
 - Functions, Logarithms and Exponentials
 - Vectors, Dot Products, Orthogonality
 - Matrices, Matrix Operations, Linear Transformations,
 - Eigendecomposition
 - Calculus, Differentiation, Integration
 - Probability and Statistics
- Python
- Numpy, Scipy – numerical/scientific computing, linear algebra
- Matplotlib – for plotting
- Scikitlearn – for machine learning

Math

- Chapter 2 of Pattern recognition and machine learning by Christopher Bishop
- Part 1 of Deep Learning book:
<http://www.deeplearningbook.org/>
- Mathematics for Machine Learning
<https://mml-book.github.io/>
- Essence of linear algebra:
<http://youtu.be/kjBOesZCoqc>
- Essence of calculus: <https://goo.gl/Hnk1jA>
- Programming
- Practice Python
- <https://try.jupyter.org/>
- <https://docs.python.org/3/tutorial/>
- Video Tutorials:
<https://www.youtube.com/watch?v=cpPG0bKHYKc>
- Play with Numpy, Matplotlib, scikitlearn

Foundations of ML : Evaluation

Tentative Evaluation scheme* :

- Exams (50%)
 - online quizzes, and Endsem (physical)
- Two Assignments (20%)
- Kaggle Project (30%)



* Evaluation scheme and weightage is tentative and is subject to change