

Manassas Machine Learning 101 and Beyond

Bridgette Townsend

October 5, 2019

Who am I?

- Bridgette Townsend
 - Software Engineer
 - Embedded Systems (vehicles, mobile devices, etc.)
 - Vulnerability Researcher
 - PhD Student
 - Data Science

Introduction

[in • tro • duc • tion]

- What is Artificial Learning/Machine Learning (ML)
- ML Fundamentals
- Getting Started (Optional)

Artificial Intelligence

[ar • ti • fi • cial • in • tell • i • gence]

Artificial Intelligence (AI) aims to create intelligent systems based on human intelligence.

Machine Learning

[ma • chine • lear • ning]

Machine learning is a subset of AI, leveraging large data sets to optimize computer systems for specific tasks.

Beginnings

[be • gi • nings]

1950: Alan Turing proposes the Turing Test as a measure of machine intelligence.

1956: AI is born at the Dartmouth conference organized by John McCarthy, father of Artificial Intelligence.

1958: John McCarthy defined the Lisp language.



Alan Turing



John McCarthy

The birth of AI - Dartmouth Conference 1956

(New Hampshire, USA)



John McCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



Nathaniel Rochester

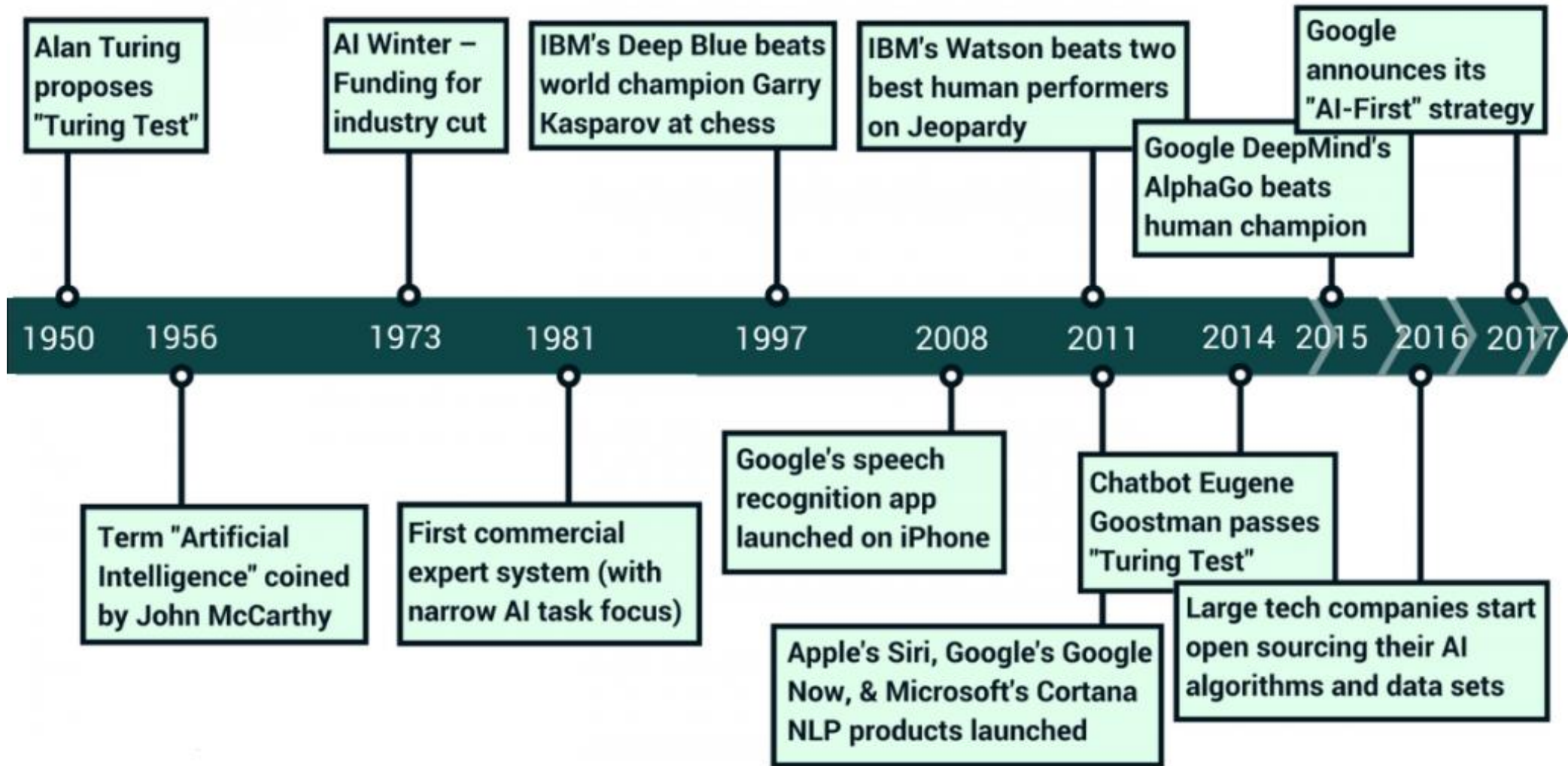


Trenchard More

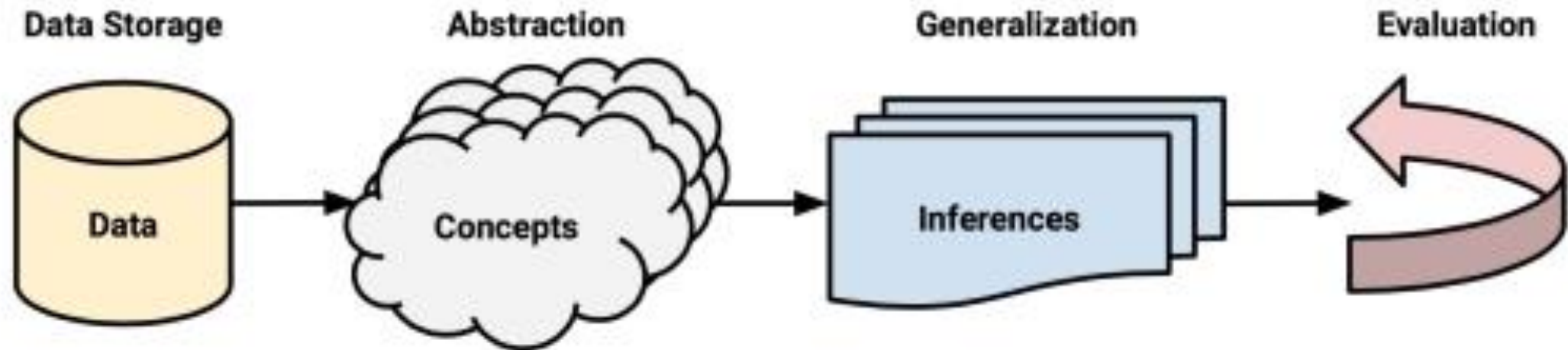
The proposal for the conference included this assertion:

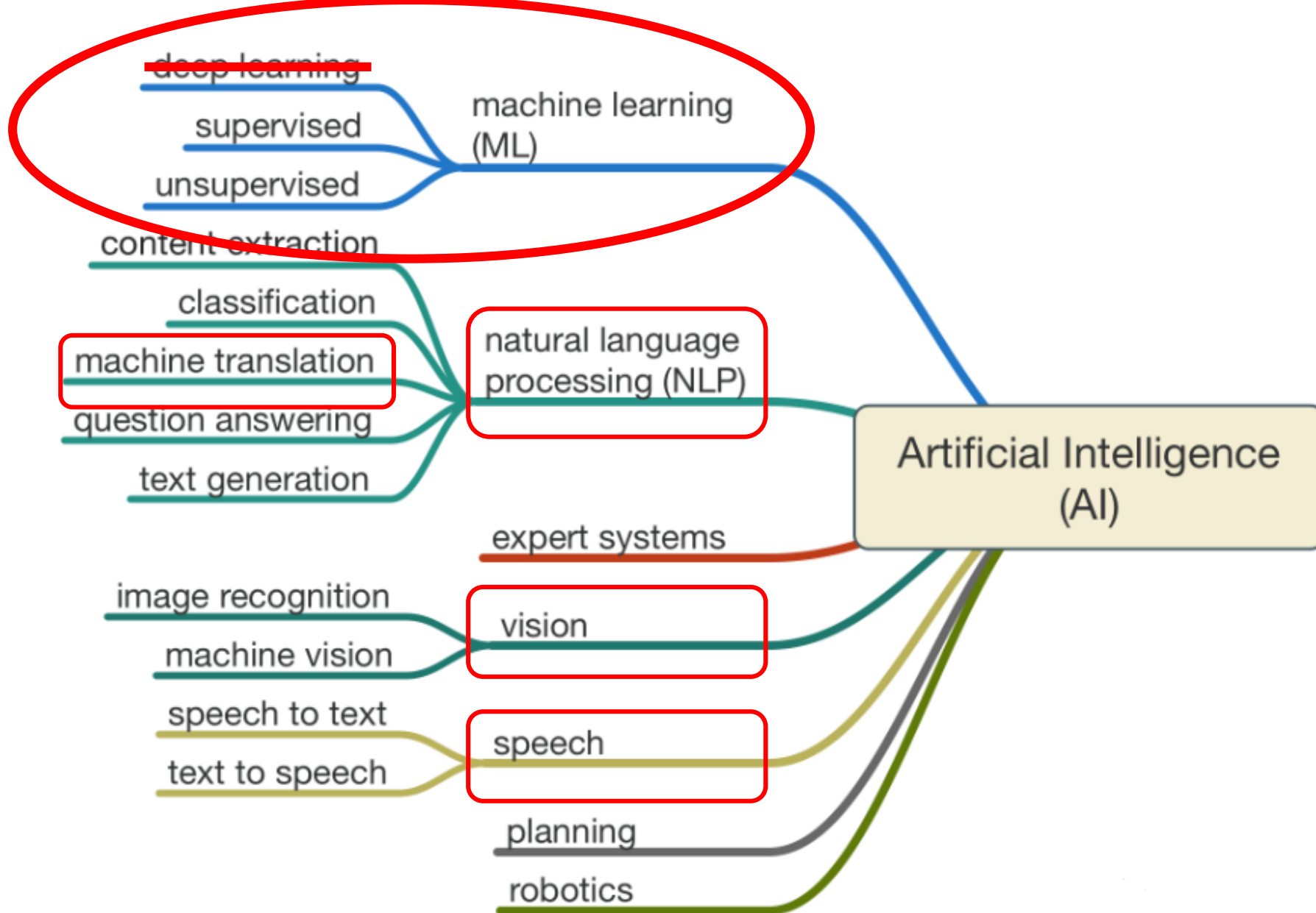
"every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it".

John McCarthy persuaded the attendees to accept "Artificial Intelligence" as the name of this new field.

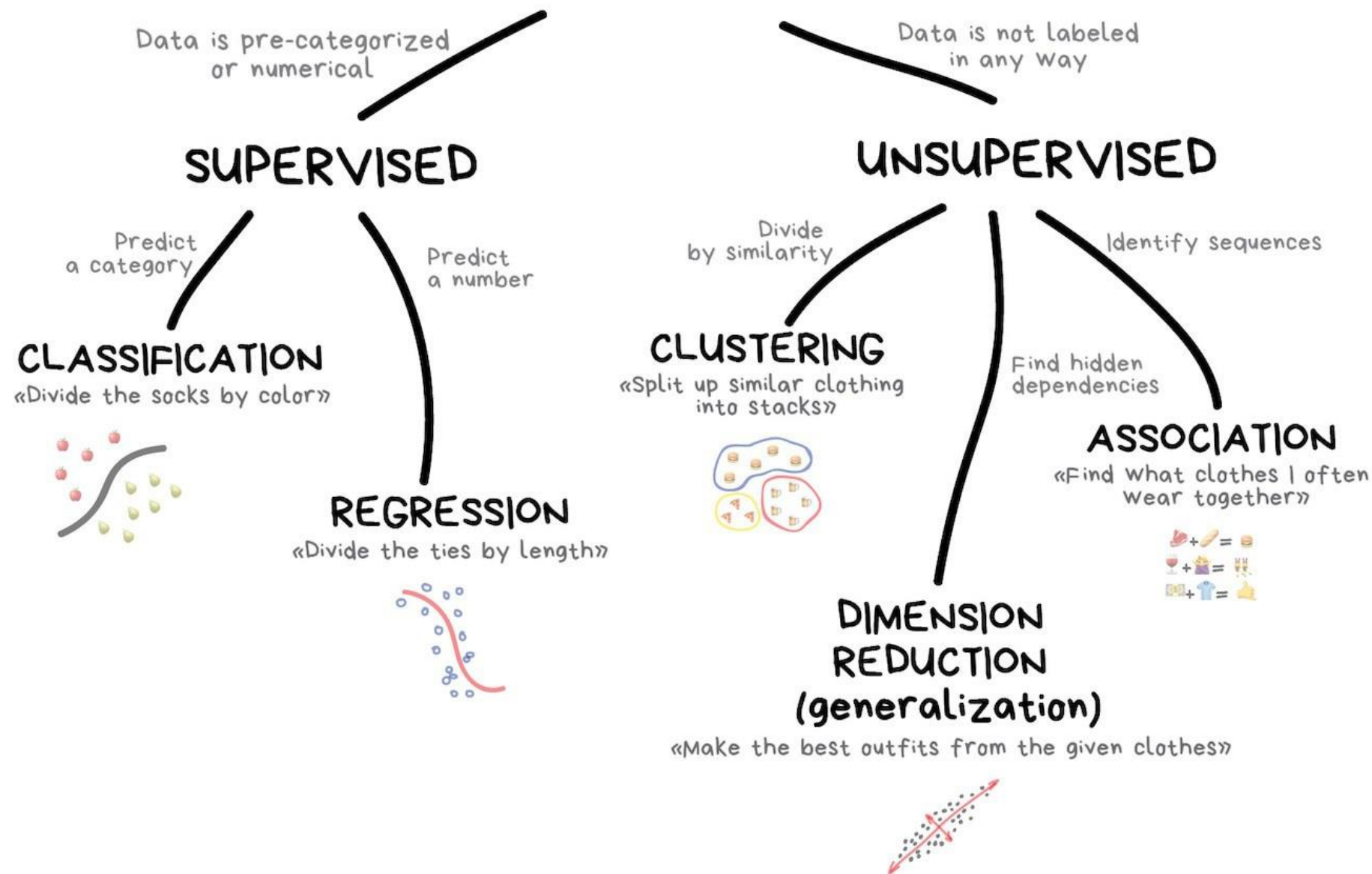


Basic Learning Process for Humans or Machines





CLASSICAL MACHINE LEARNING



Supervised Learning Algorithms

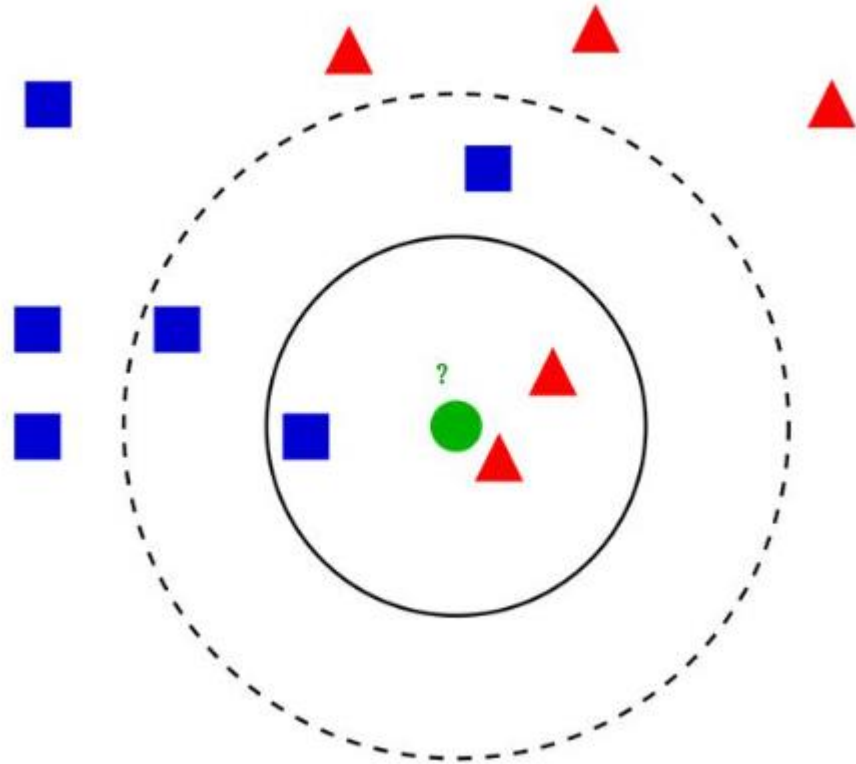
- K Nearest Neighbors
- Naïve Bayes Classifier
- Regression Methods
- Support Vector Methods
- Decision Trees

Unsupervised Learning Algorithms

- Clustering
 - K Means Clustering
- Self-Organizing Maps

K nearest neighbors (KNN)

- Supervised Learning technique
- Take classified data represented by a vector of features
- Measure distances between the new data point and the nearest K of our already-classified data and conclude to which class our new data point belongs.
- Can also be used for regression, instead of having classes we can associate a number with each data point



Naïve Bayes Classifier

- Supervised Learning
- Samples of data representative of different classes
- Calculate the probability of new data being in each class

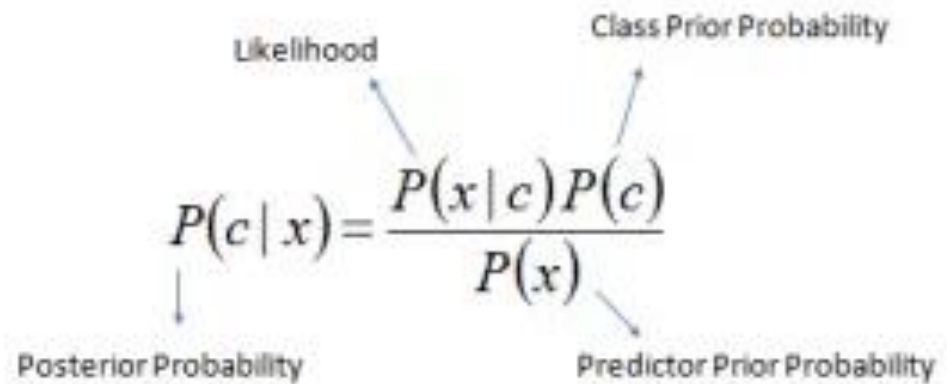
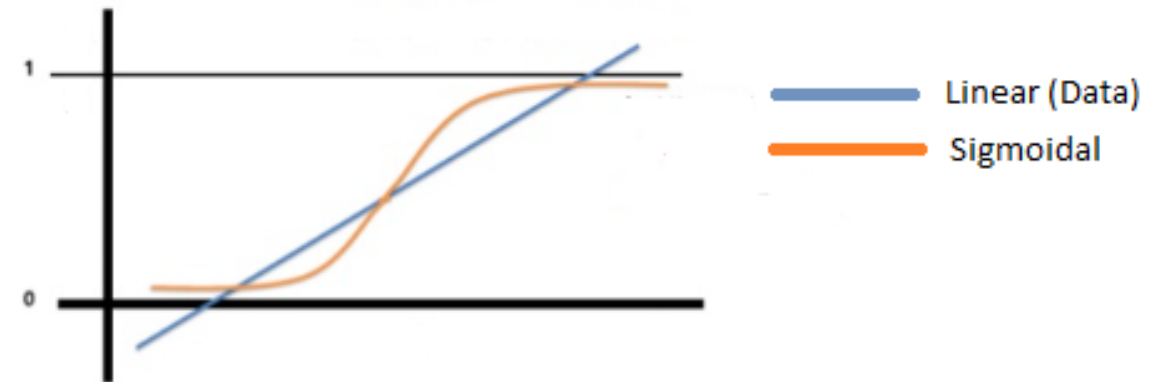
$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$


Diagram illustrating the components of the Naïve Bayes formula:

- $P(c|x)$ is labeled as Posterior Probability.
- $P(x|c)$ is labeled as Likelihood.
- $P(c)$ is labeled as Class Prior Probability.
- $P(x)$ is labeled as Predictor Prior Probability.

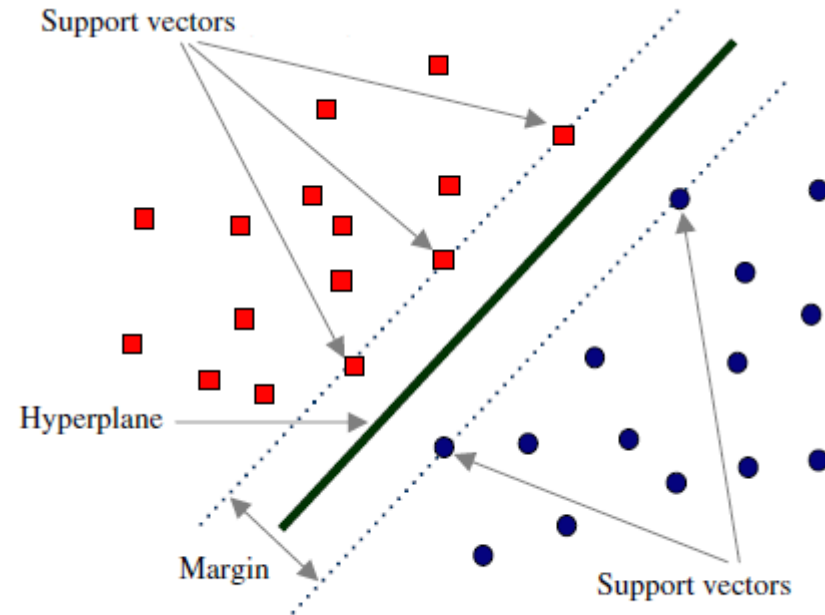
Regression Methods

- Supervised Learning
- They try to explain a dependent variable in terms of independent variables
- Independent variables are numerical
- Fit straight-line, polynomials or other functions to predict the dependent variables
- Can also be used for classification when the dependent variables are usually zeros and ones



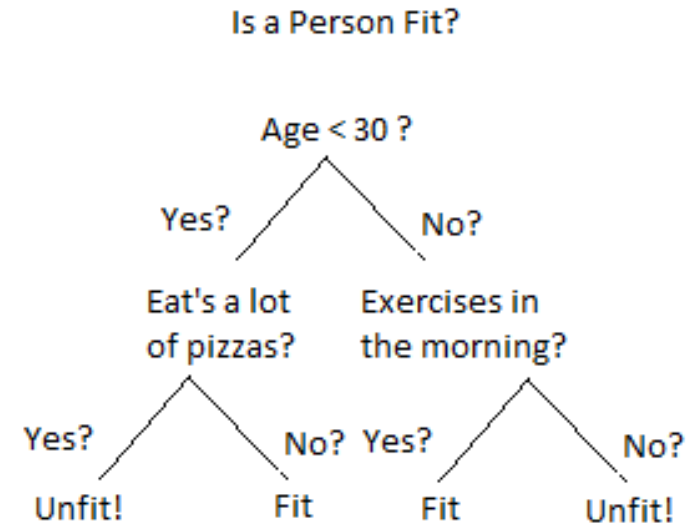
Support Vector Machines (SVM)

- Supervised Learning
- Used for classification
- Classified data represented by vectors of features
- Divides data according to which side of a hyperplane in feature space each point lies



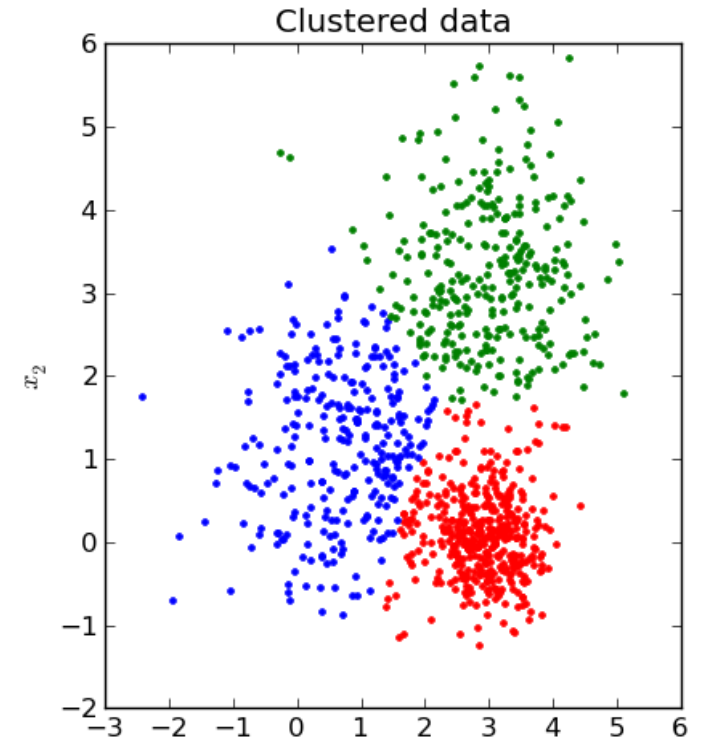
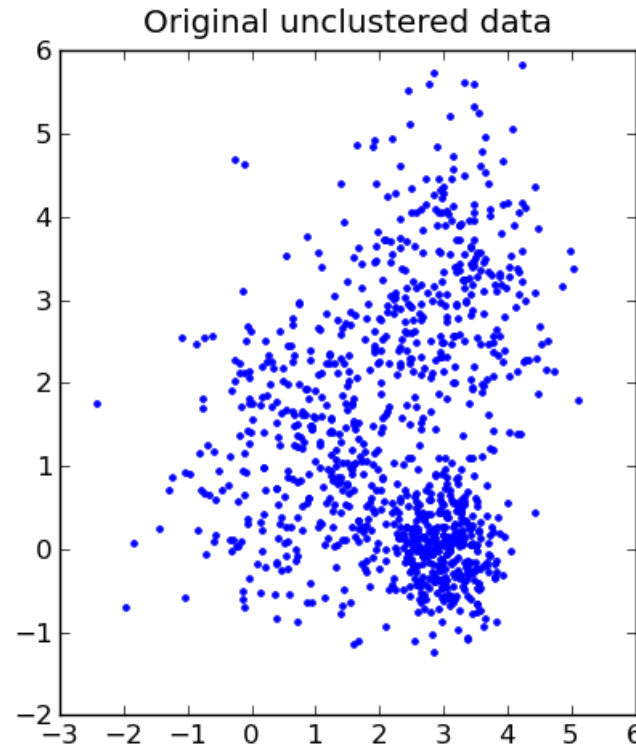
Decision Trees

- Supervised learning
- Can be used for classification and regression
- A decision tree is just a flowchart



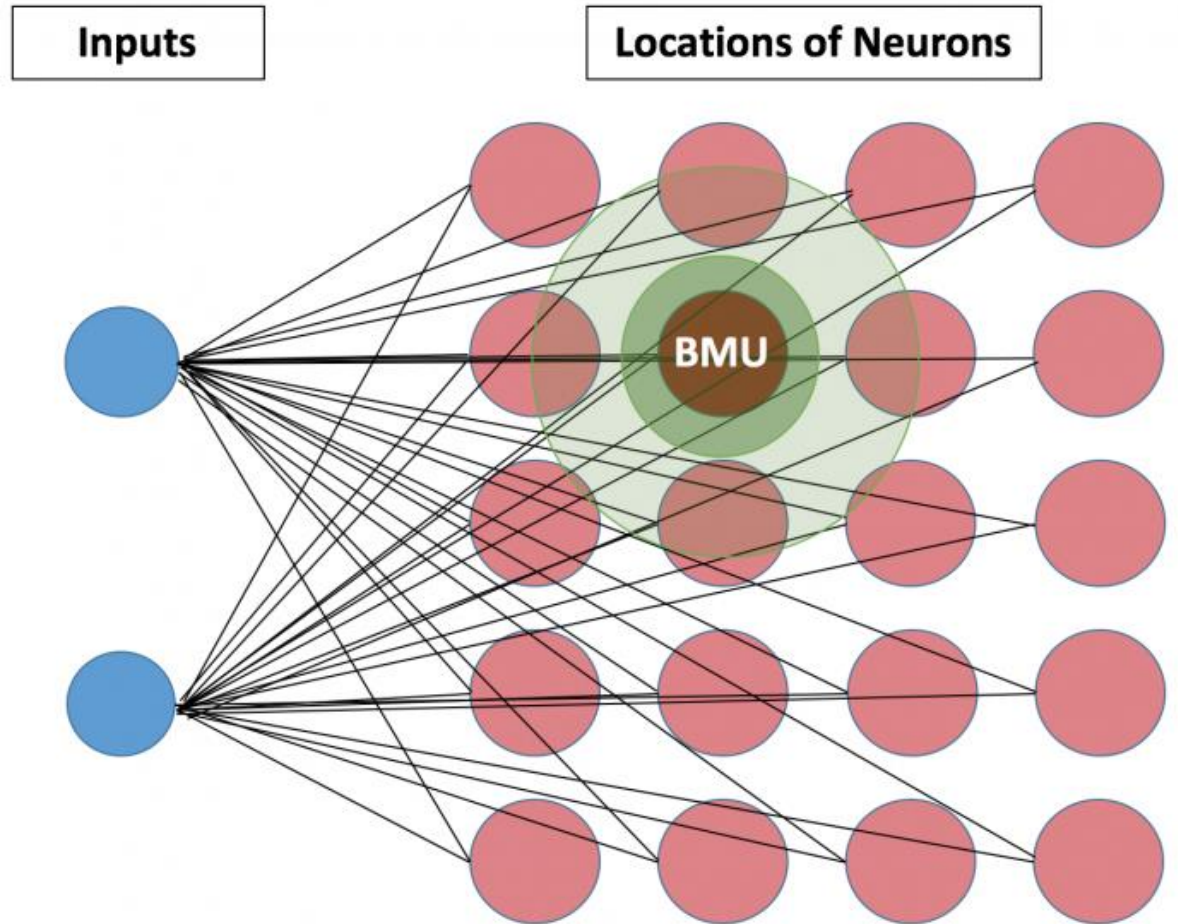
K Means Clustering

- Unsupervised Learning technique
- Each data point is represented as a vector
- Each entry in the vector represents a feature
- Data points are not labeled or classified
- Data points are grouped in a sensible way



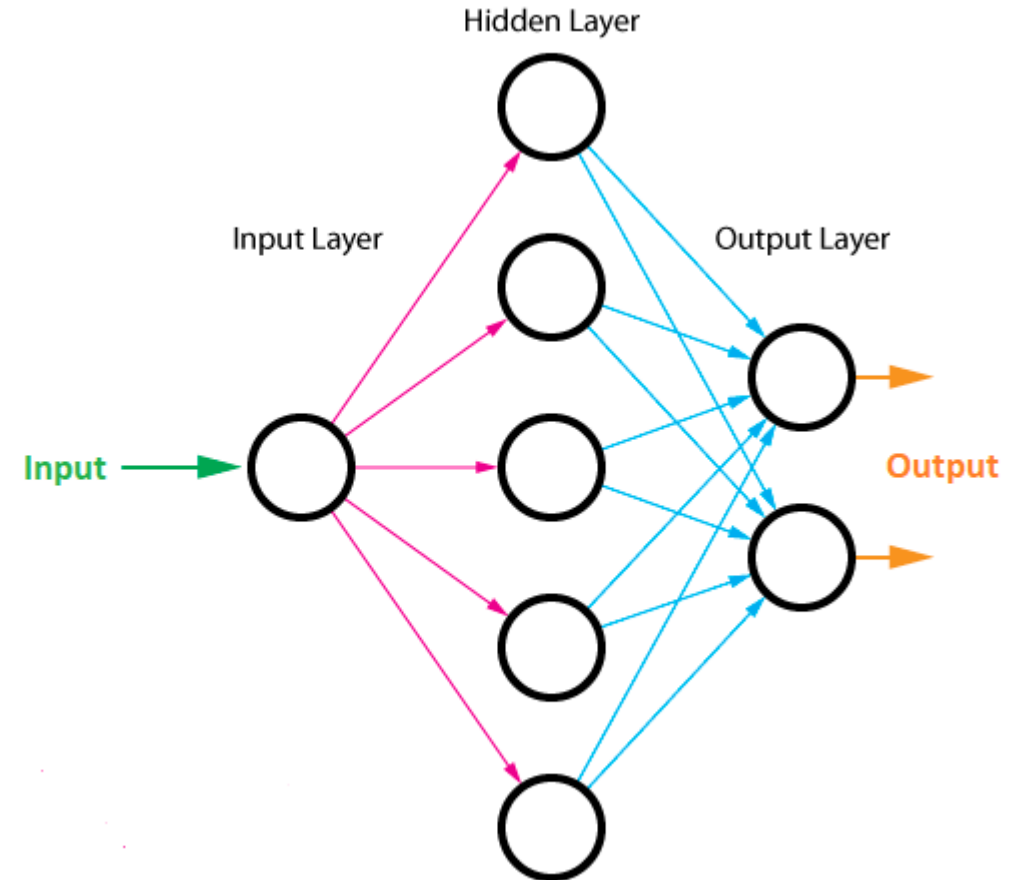
Self-Organizing Maps

- Unsupervised Learning
- Start off with vectors of features for all data points
- They are grouped together according to how similar the vectors are
- Some people think of this technique as a version of a neural network
 - The method is a typically 2-d picture consisting of square array of cells.
 - A particular kind of SOM known as a kohonen networks. This SOM has a feed-forward structure with a single computational layer arranged in row and columns.



Neural Networks

- Can be used for supervised or unsupervised learning
- Meant to mimic what happens in the brain
- An input vector is transformed typically by multiplying a matrix
- Deep Learning – neural nets with deep layers. Increased number of hidden layers and nodes that integrate to estimate the output(s)



- Predictive maintenance or condition monitoring
- Warranty reserve estimation
- Propensity to buy
- Demand forecasting
- Process optimization
- Telematics

Manufacturing



- Predictive inventory planning
- Recommendation engines
- Upsell and cross-channel marketing
- Market segmentation and targeting
- Customer ROI and lifetime value

Retail



- Alerts and diagnostics from real-time patient data
- Disease identification and risk stratification
- Patient triage optimization
- Proactive health management
- Healthcare provider sentiment analysis

Healthcare and Life Sciences



- Aircraft scheduling
- Dynamic pricing
- Social media – consumer feedback and interaction analysis
- Customer complaint resolution
- Traffic patterns and congestion management

Travel and Hospitality



- Risk analytics and regulation
- Customer Segmentation
- Cross-selling and up-selling
- Sales and marketing campaign management
- Credit worthiness evaluation

Financial Services



- Power usage analytics
- Seismic data processing
- Carbon emissions and trading
- Customer-specific pricing
- Smart grid management
- Energy demand and supply optimization

Energy, Feedstock, and Utilities

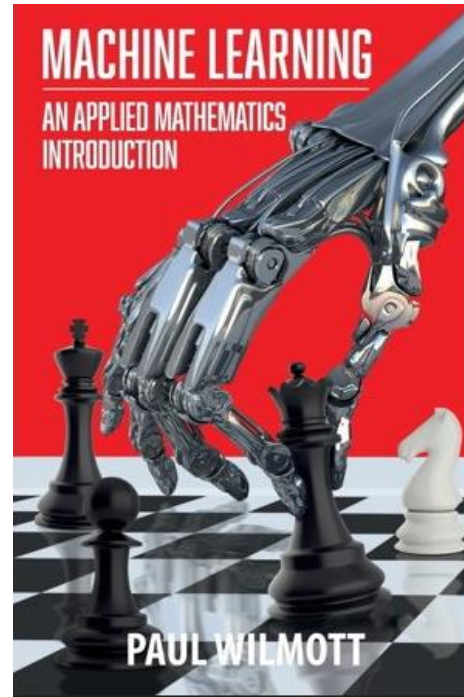


Getting Started

- Tons of information available on YouTube
- Take Andrew Ng's machine learning course
- Download Jupyter Notebook or Anaconda Python/R Distribution to start programming
- Read books, papers and blogs
- Keep coming to this meetup 😊

Book References on material covered today

- Machine Learning: An Applied Mathematics Introduction by Paul Wilmott
- Machine Learning by Tom M. Mitchell



Recommended Papers

- Computing Machinery And Intelligence by Alan M. Turing
- Mastering The Game of Go without Human Knowledge by David Silver, et. al.