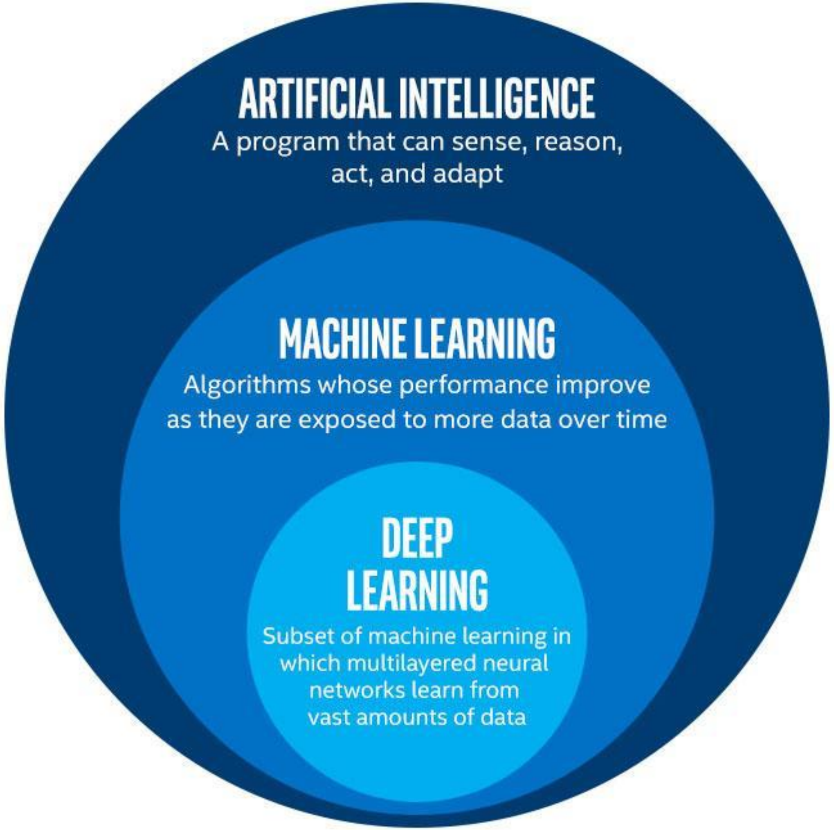
Introduction to ML

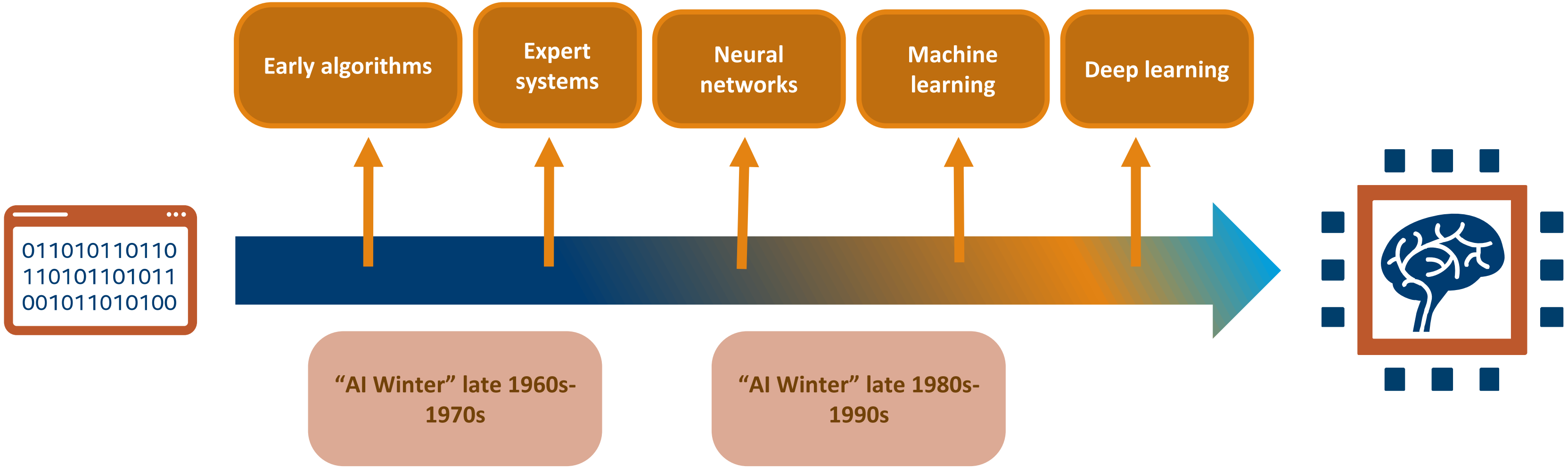
Definitions

- **Artificial Intelligence**
  - Human Intelligence Exhibited by Machines
- **Machine Learning**
  - An Approach to Achieve Artificial Intelligence
- **Deep Learning**
  - A Technique for Implementing Machine Learning



History of AI

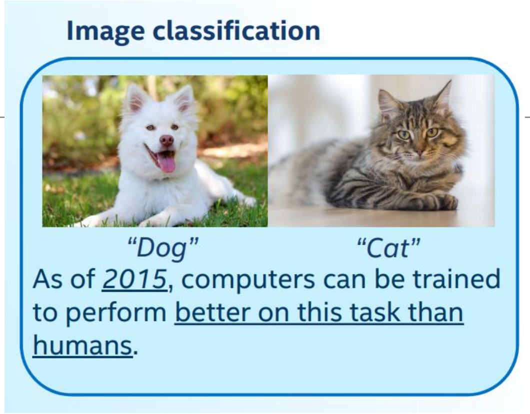
AI has experienced several hype cycles, where it has oscillated between periods of excitement and disappointment.

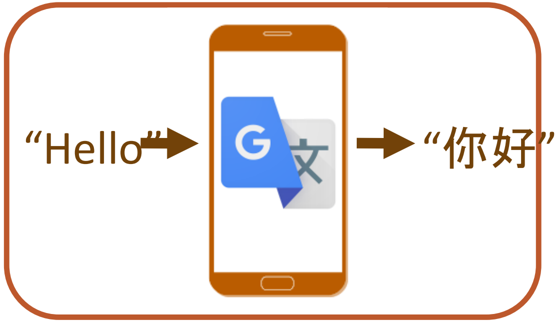


Modern AI

Deep Learning Breakthroughs (2012 – Present)

- In 2012, deep learning beats previous benchmark on the ImageNet competition.
- In 2013, deep learning is used to understand "conceptual meaning" of words.
- In 2014, similar breakthroughs appeared in language translation.
- These have led to advancements in Web Search, Document Search, Document Summarization, and Machine Translation.

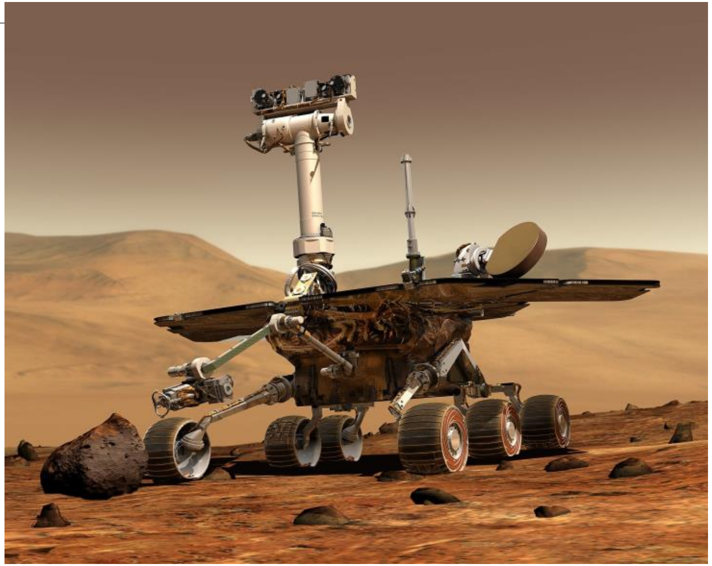




Google Translate

Deep Learning Breakthroughs (2012 – Present)

- In 2014, computer vision algorithm can describe photos.
- In 2015, Deep learning platform TensorFlow is developed.
- In 2016, DeepMind's AlphaGo, developed by Aja Huang, beats Go master Lee Se-dol.



Autonomous Mars rover

Modern AI (2012 – Present): Deep Learning Impact

Computer vision

Self-driving cars: object detection

Healthcare: improved diagnosis

Natural language

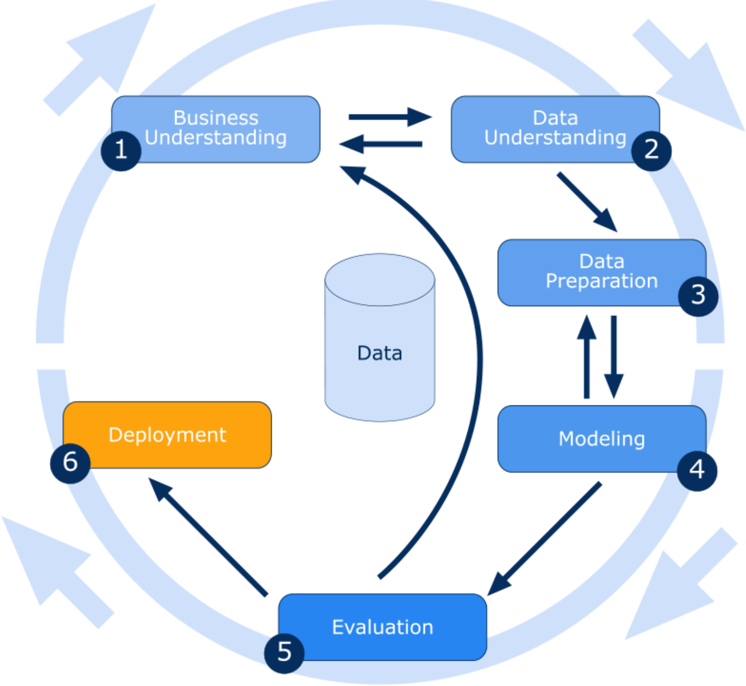
Communication: language translation

AI Applications

AI Usage Growth

Data Science Workflow

CRiSP-DM (Data Science work flow)



The cross industry standard process for data mining, CRISP-DM, is a data mining process model that data mining experts use to tackle problems

Very iterative process

Approach for Data Science Projects

Business Problem Definition

- Define business objectives and frame the problem
- Layout project plans
- Validate with stake-holders

Through the Initial discussion with stake holders

Understand SYSTEMS, Data and processes

- Understand systems involved and business process
- Gain domain knowledge
- Collect, describe and explore the available data
- Verify data quality

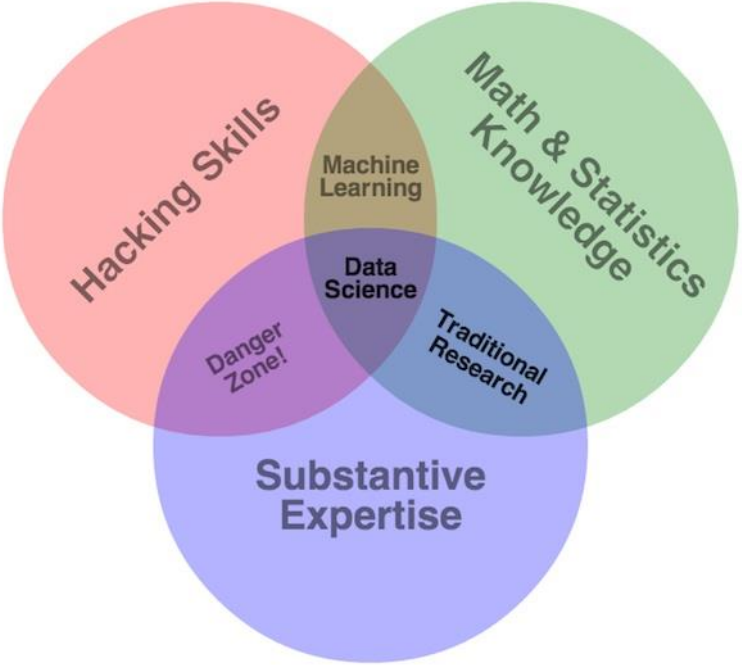
Evaluate and choose Solution approach

- What are solution paths to address the need.
- Select simplest and proven solution path - consult with experts that have solved similar problems in the domain

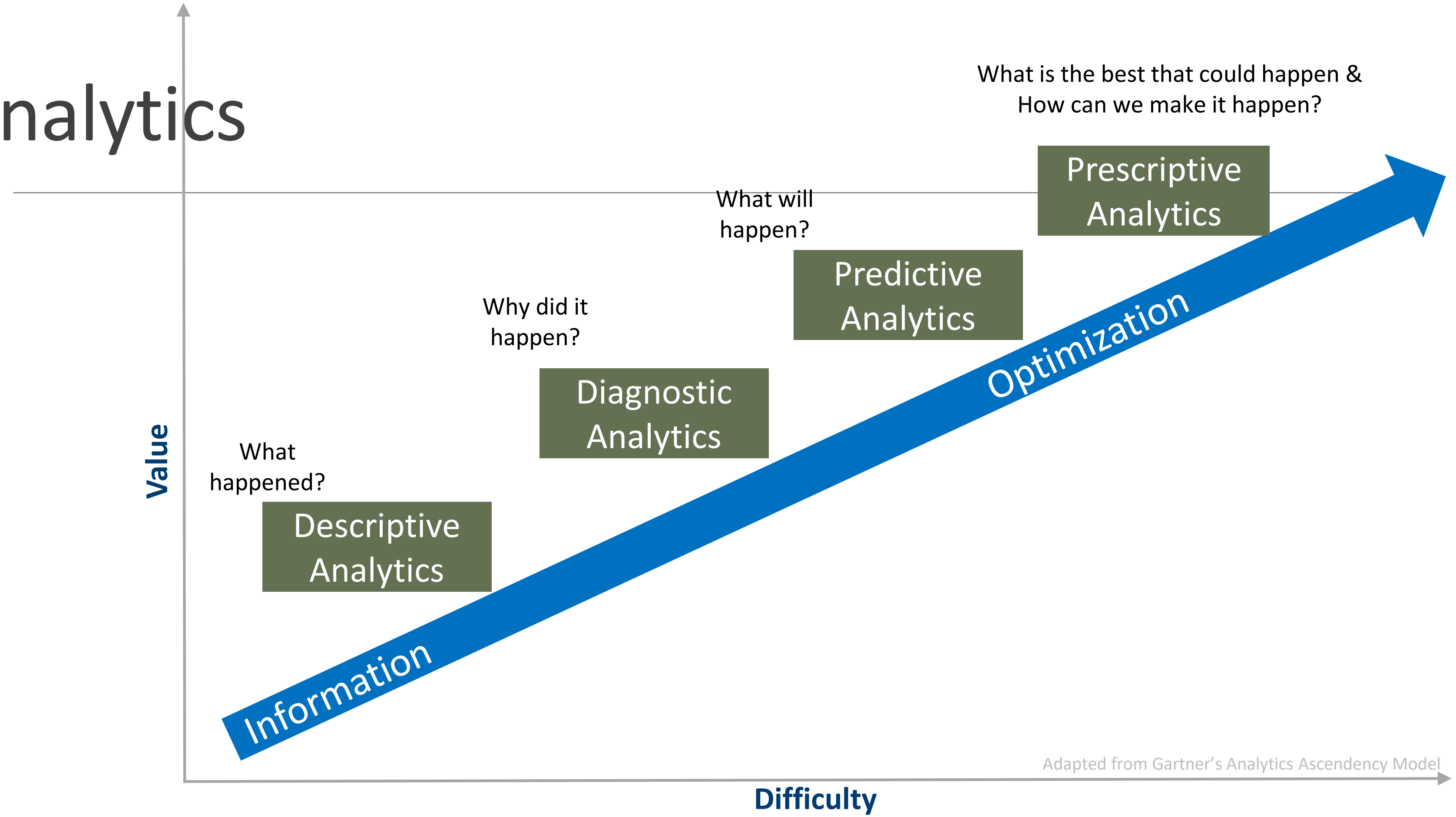
Implement, Test, Validate

- Develop/implement chosen algorithm, validate
- Pilot & ensure the solution properly addresses business problem
- Produce deployment plan that meets sustainability needs

Data Science Skill Sets



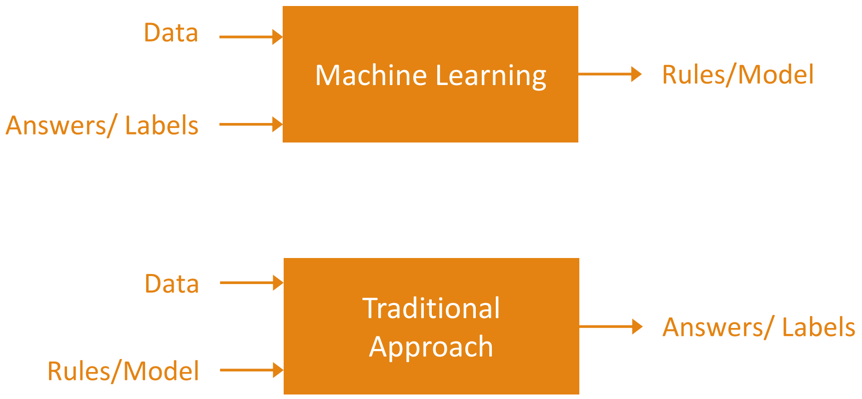
Analytics


What is the best that could happen &
How can we make it happen?
Prescriptive Analytics
What will happen?
Predictive Analytics
Why did it happen?
Diagnostic Analytics
What happened?
Descriptive Analytics
Value
Information
Optimization
Difficulty
Adapted from Gartner's Analytics Ascendency Model

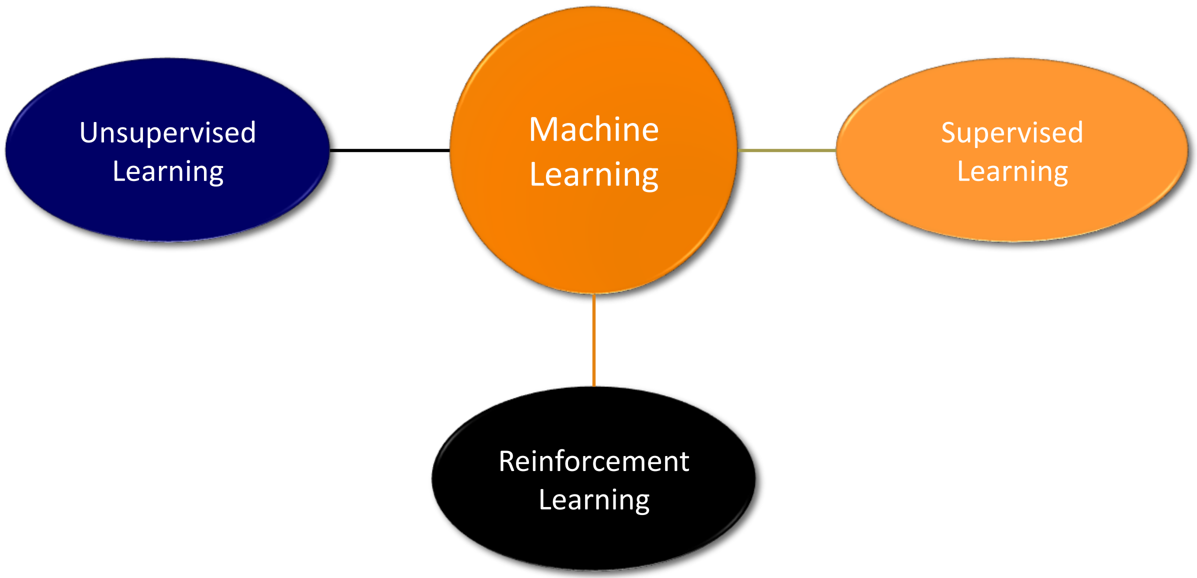
Machine Learning

Machine Learning

Machine Learning relates with the study, design, and development of models and algorithms that give computers the capability to learn from data, instead of requiring explicit programming of hard-coded rules/logic.



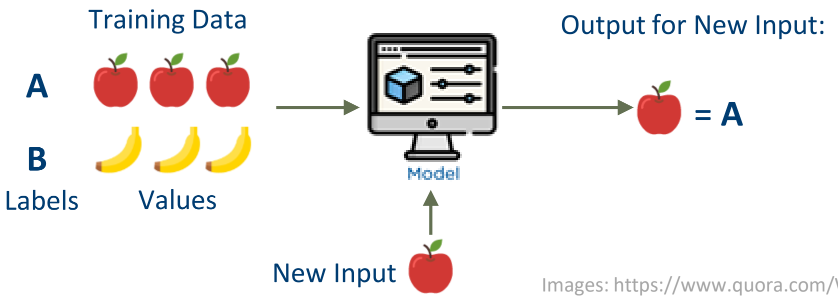
Machine Learning algorithms



Supervised vs. unsupervised learning

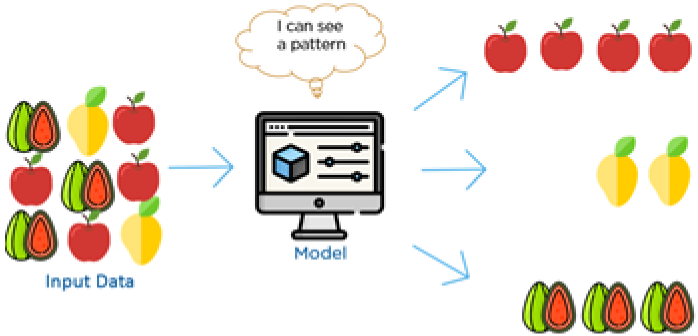
Supervised

- Begin with data we **already know the desired output**
- Use **Training data** to derive relationships between the input features and desired output
- Predict future outcomes from derived relationships from training data



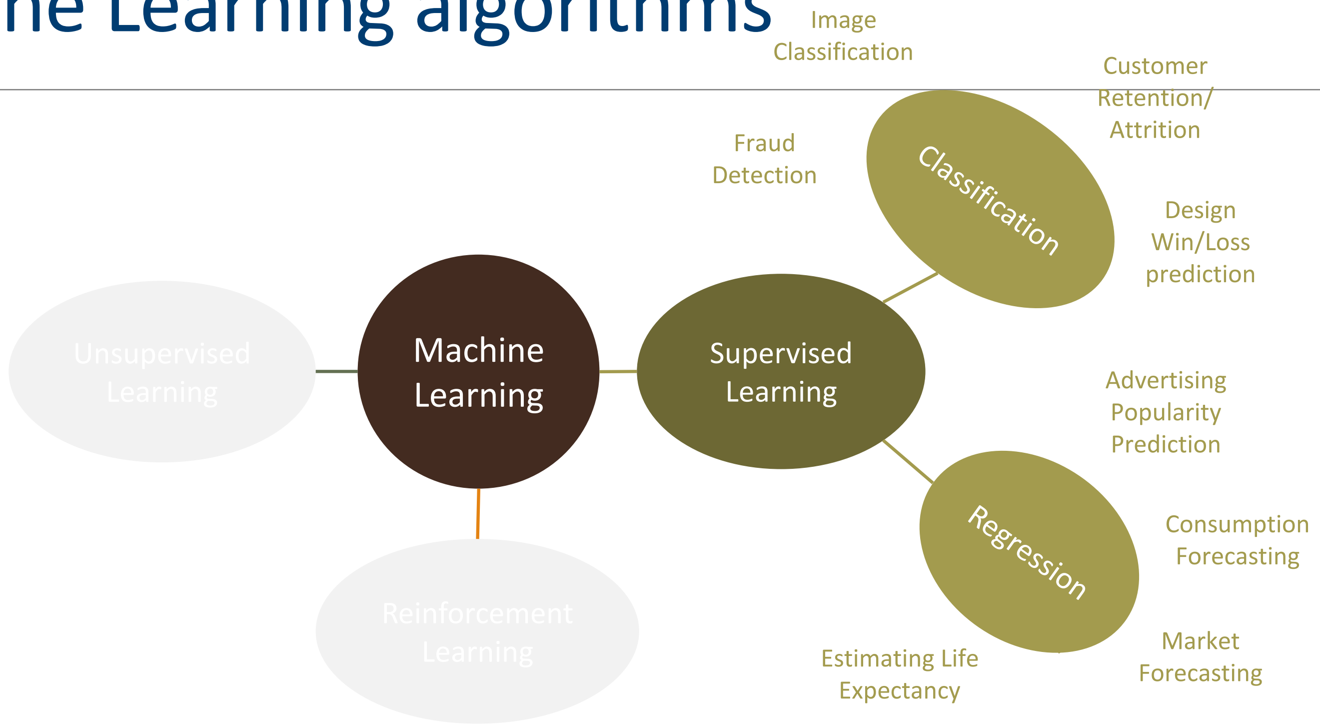
Unsupervised

- Begin with little or no idea what our results should look like
- Derive **structure** from data where we don't necessarily know the effect of the variables.



Images: https://www.quora.com/What-is-the-difference-between-supervised-and-unsupervised-learning-algorithms

Machine Learning algorithms

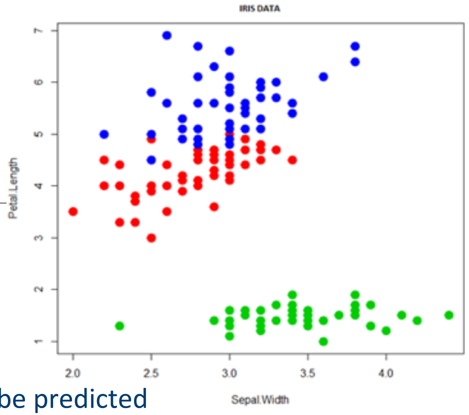


Supervised Learning – Classification

Iris setosa | Iris versicolor | Iris virginica

Input / observation / attribute: categorical or numeric

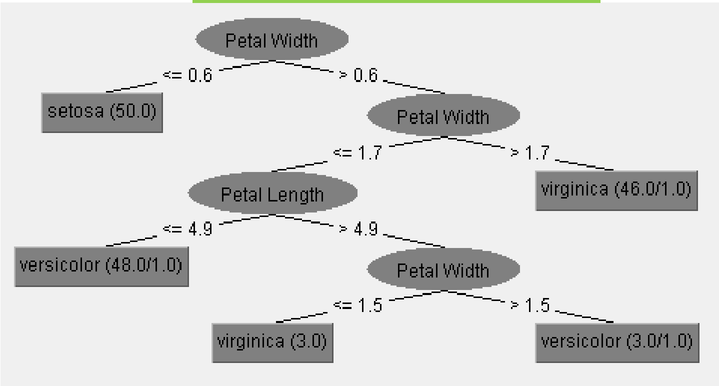
Output / response / label: categorical



Target: Columns to be predicted

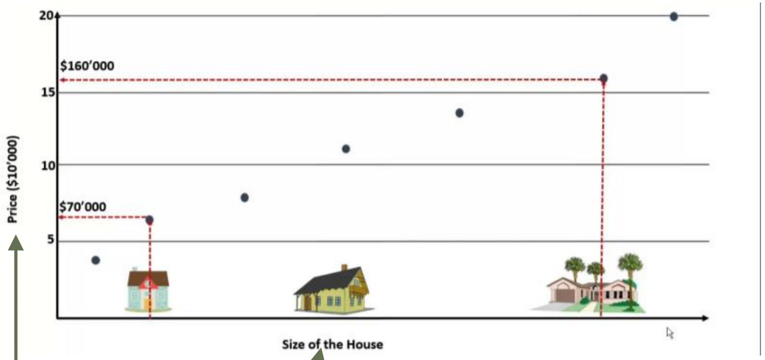
sepal length	sepal width	petal length	petal width	Flowername
5.1	3.5	1.4	0.2	Iris-setosa
4.9	3	1.4	0.2	Iris-setosa
4.7	3.2	1.3	0.2	Iris-setosa
7	3.2	4.7	1.4	Iris-versicolor
6.4	3.2	4.5	1.5	Iris-versicolor
6.9	3.1	4.9	1.5	Iris-versicolor
6	3	4.8	1.8	Iris-virginica
6.9	3.1	5.4	2.1	Iris-virginica
6.7	3.1	5.6	2.4	Iris-virginica

Decision Tree - Algorithm



Features : Attributes of the Data

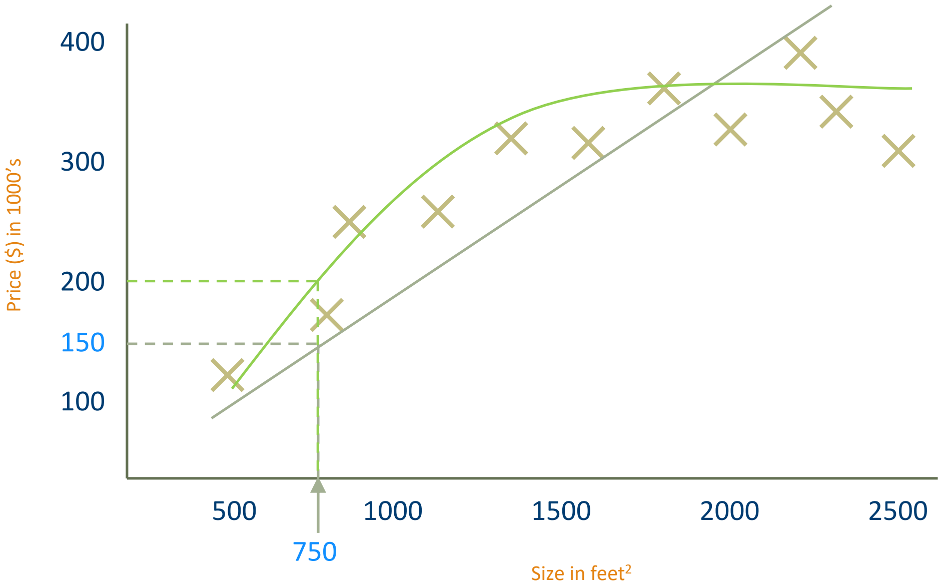
Supervised Learning – linear Regression Example



Label/Dependent variable: Price of the house in Dollars

Features : Size of the House in Square Feet's(Independent Variable)

House price prediction



Evaluation Metric

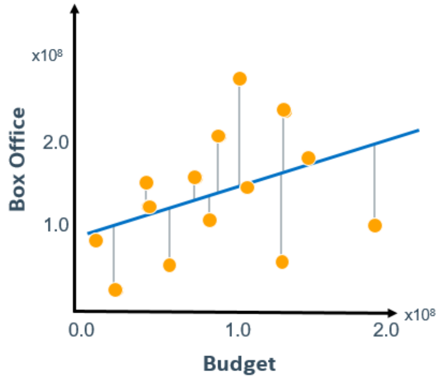
There are many metrics available* to measure performance, such as:

Accuracy target

- **Accuracy**: how well predictions match true values.

- **Mean Squared Error**: average square distance between prediction and true value.

$$\min_{\beta_0,\beta_1} \frac{1}{m} \sum_{i=1}^{m} \left(\left(\beta_0 + \beta_1 x_{obs}^{(i)} \right) - y_{obs}^{(i)} \right)^2$$



**The wrong metric can be misleading or not capture the real problem.*

Which Model?

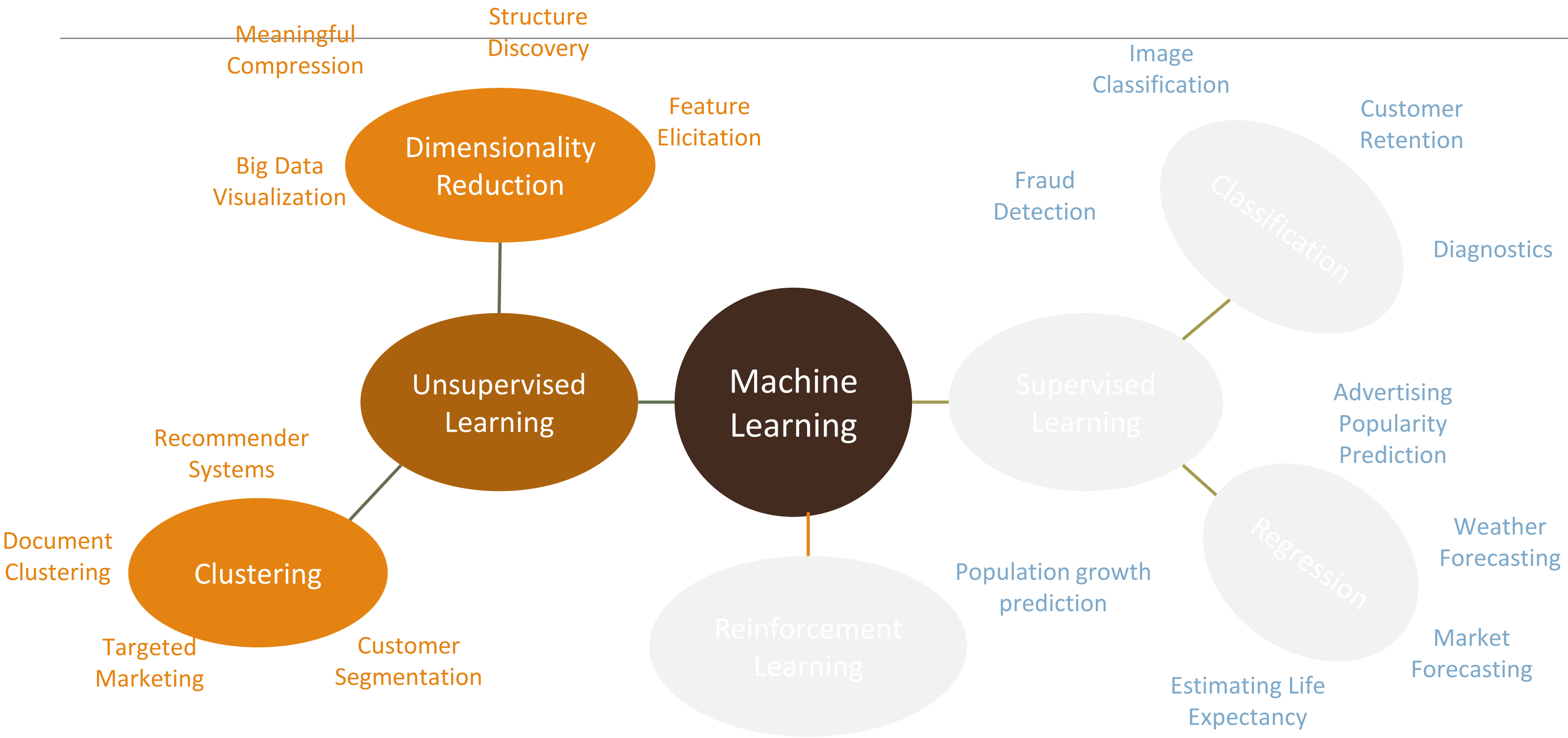
Some considerations when choosing are:

- Time needed for training
- Speed in making predictions
- Amount of data needed
- Type of data
- Problem complexity
- Ability to solve a complex problem
- Tendency to overcomplicate a simple one

Supervised Learning Algorithms

Algorithm	Classification/Regression	Comments
K-Nearest Neighbors	Classification & Regression	Instance based, suffers from curse of dimensionality
Linear Regression	Regression	Need follow certain assumptions
Naïve Bayes	Classification	Probability based. Inputs need to independent from each other.
Logistic Regression	Classification	Need follow certain assumptions
Support Vector Machines	Classification & Regression	Where class boundaries are well separated.
Decision Trees	Classification & Regression	Tree based. Not effective with low covariance.
Neural Networks	Classification & Regression	Non-linear functional approximation. Also used for unsupervised learning (DBNs)
Ensemble (Random Forests, Boosting, Bagging, Bucket of models, Stacking, etc)	Classification & Regression	Improved prediction performance from multiple hypothesis.

Machine Learning algorithms



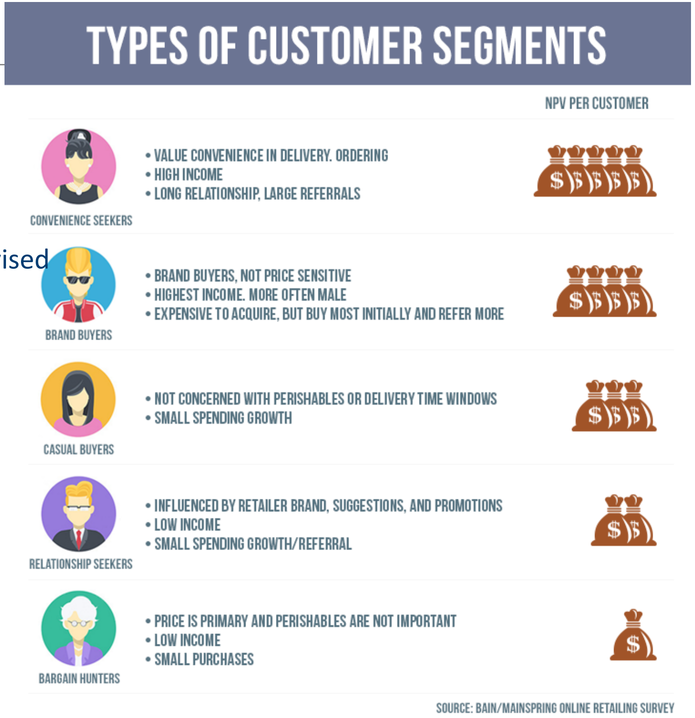
Un Supervised – Customer Segmentation

All Customers data from previous Transactions

Few example features Extracted

1. Since how long been with us
2. #of times visits per month
3. Customer Gender and age
4. #of Referrals generated
   .....

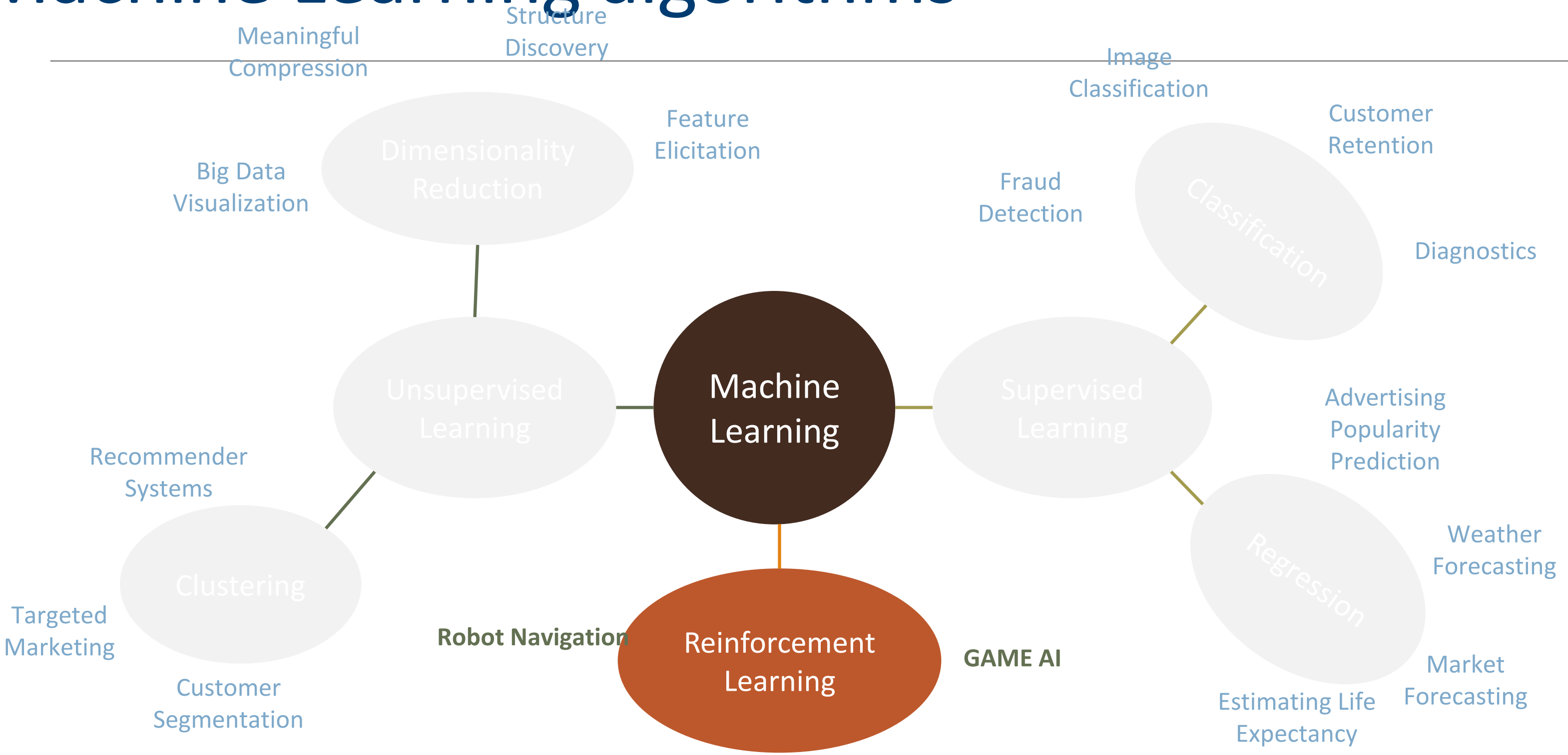
Apply Unsupervised Algorithm



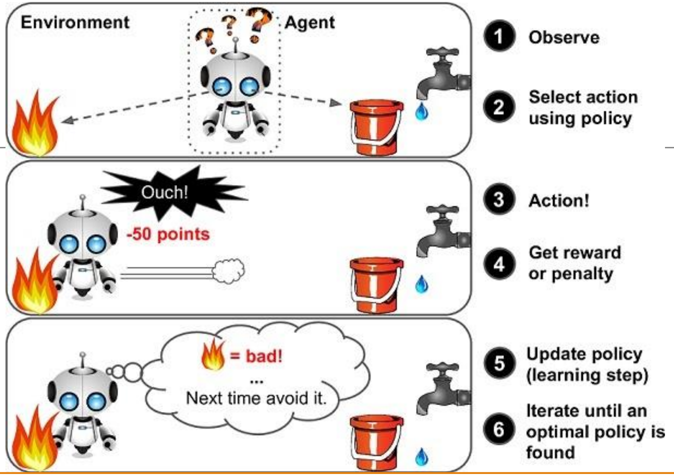
Un Supervised – Market Basket Analysis

Market Basket Analysis

Machine Learning algorithms



Elements of Reinforcement Learning



Agent: A robot / system/ a game player - capable of taking action

Environment: The universe the agent interacts with. Could be rules of a game for game playing, or other objects on road for a self driving car.

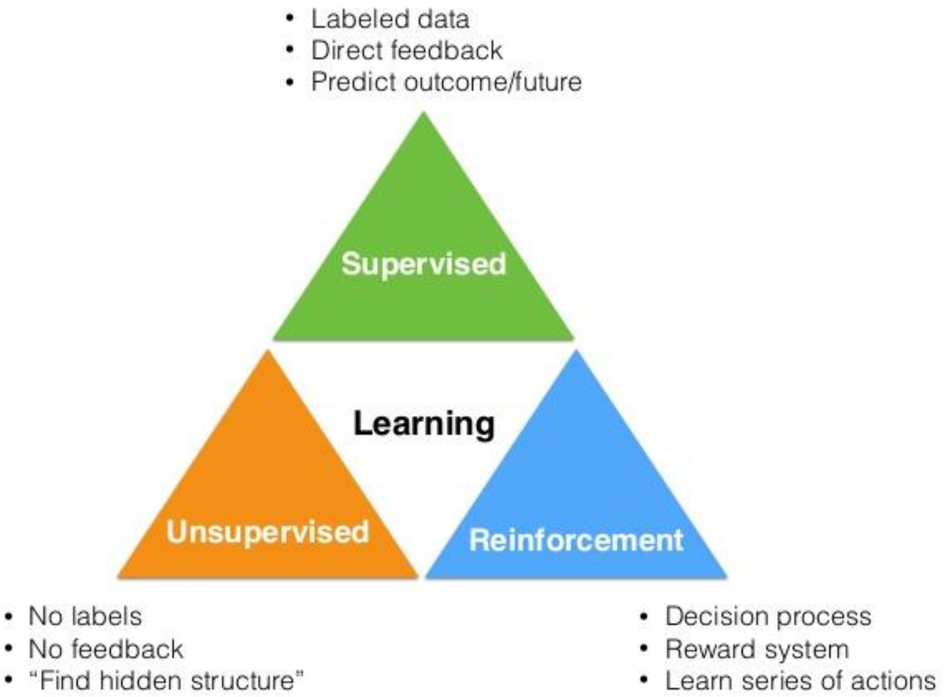
Policy: Defines the learning agent's way of behaving at a given time & environmental conditions.

Reward signal: Defines the goal in a reinforcement learning problem. On each time step, the environment sends to the reinforcement learning agent a single number, a *reward*. The agent's sole objective is to maximize the total reward it receives over the long run.

Value function: Specifies what is good in the long run. The value of a state is the total amount of reward an agent can expect to accumulate over the future, starting from that state. Whereas rewards determine the immediate, intrinsic desirability of environmental states

Reinforcement learning algorithms try to find the best ways to earn the greatest reward. Rewards can be winning a game, earning more money or beating other opponents. They present state-of-art results on very human task
In a nutshell, it is used to determine how to navigate uncertain environments in the most effective way.

Quick r



Algorithm learns to react to an environment

Source: https://www.saagie.com/blog/machine-learning-pour-les-grand-meres

Machine Learning in our Daily Lives

SPAM FILTERING	WEB SEARCH	POSTAL MAIL ROUTING
FRAUD DETECTION	MOVIE RECOMMENDATIONS	VEHICLE DRIVER ASSISTANCE
WEB ADVERTISEMENTS	SOCIAL NETWORKS	SPEECH RECOGNITION

Training and Testing Data

Using the dataset for training

Training Set: Data used during the training process.

Test Set: Data used to measure performance, simulating unseen data*.

sepal length	sepal width	petal length	petal width	species
6.7	3.0	5.2	2.3	virginica
6.4	2.8	5.6	2.1	virginica
4.6	3.4	1.4	0.3	setosa
6.9	3.1	4.9	1.5	versicolor
4.4	2.9	1.4	0.2	setosa
4.8	3.0	1.4	0.1	setosa
5.9	3.0	5.1	1.8	virginica
5.4	3.9	1.3	0.4	setosa
4.9	3.0	1.4	0.2	setosa
5.4	3.4	1.7	0.2	setosa

Training Set

Testing Set

**Not used during the training process.*

Curve Fitting Problem(Over & Under fit)

Problem: Unseen data isn't available during training.

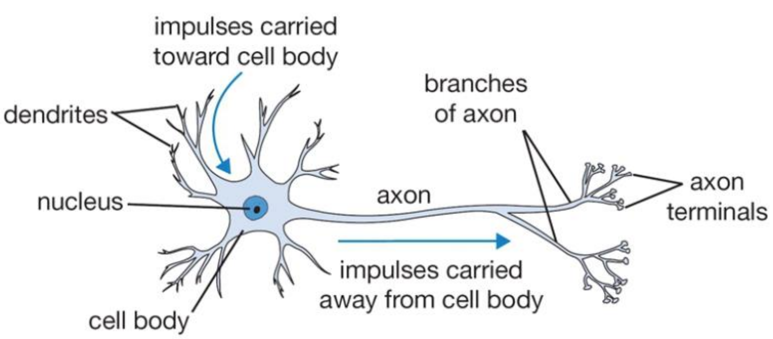
- How can performance be estimated?

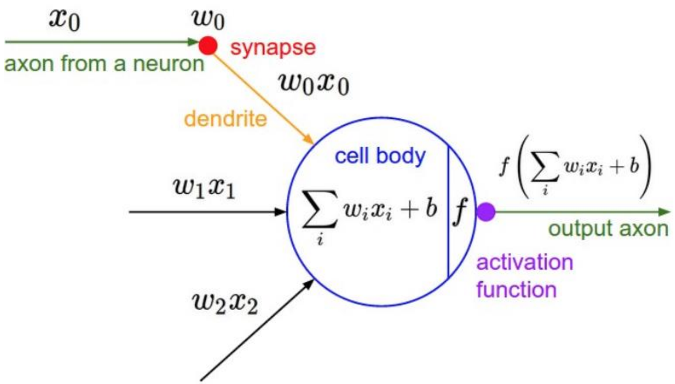
When measuring performance on the training data, there is a tendency to overfit.

Deep Learning

Neural Networks

Neural Networks: Algorithms that try to mimic Human brain, was widely used in 80's and early 90's and popularity diminished in late 90's. Recent resurgence due to availability of Compute and Data.

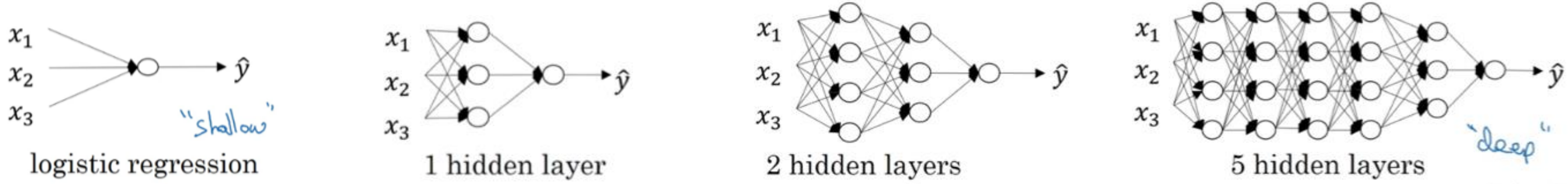




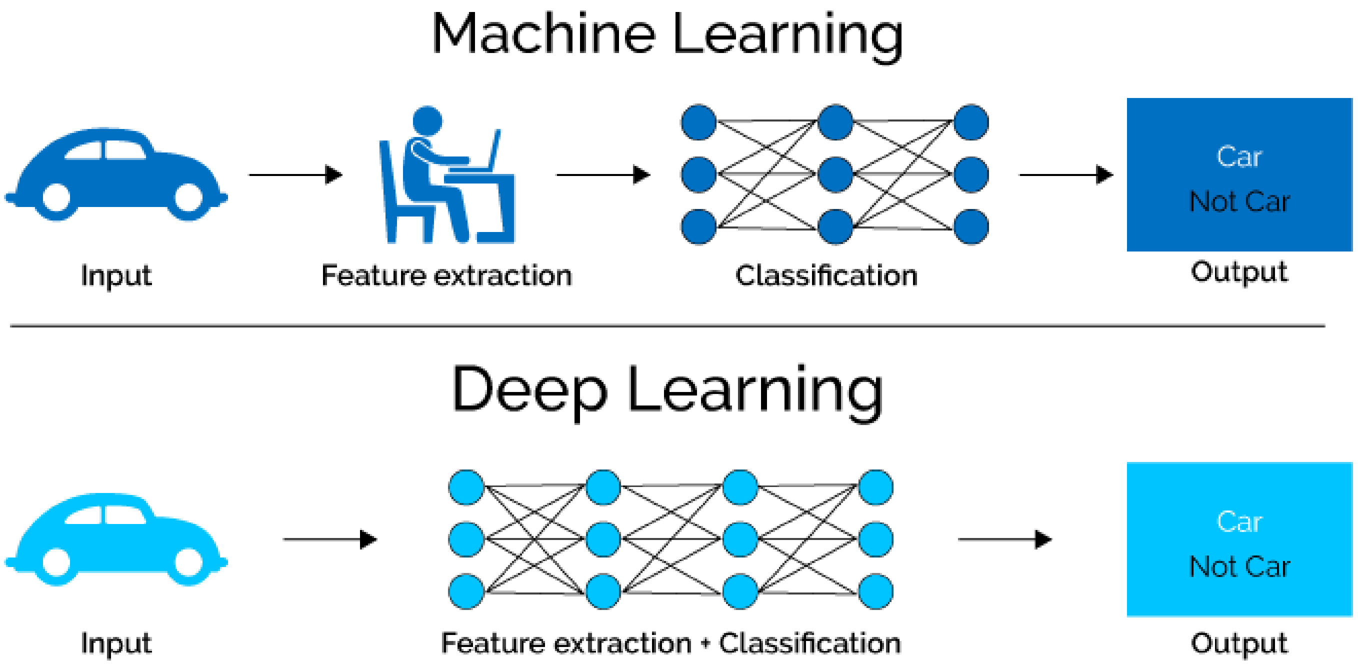
DEEP LEARNING

Deep Learning, refers to training Neural Networks, sometimes very large Neural Networks.

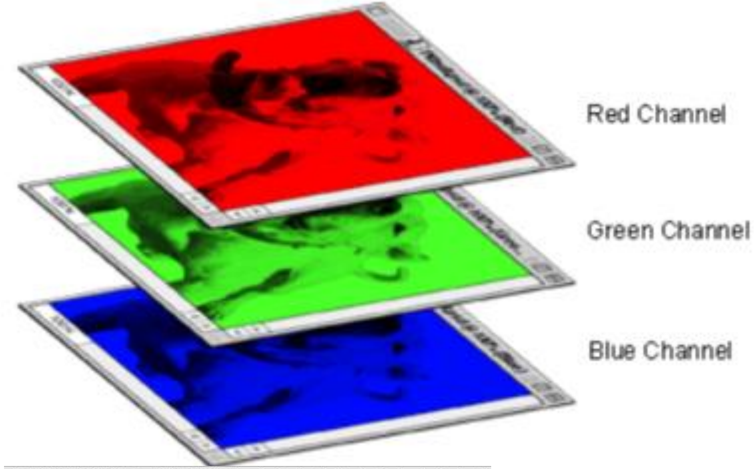
What is Deep Neural Network?



Why is Deep learning required



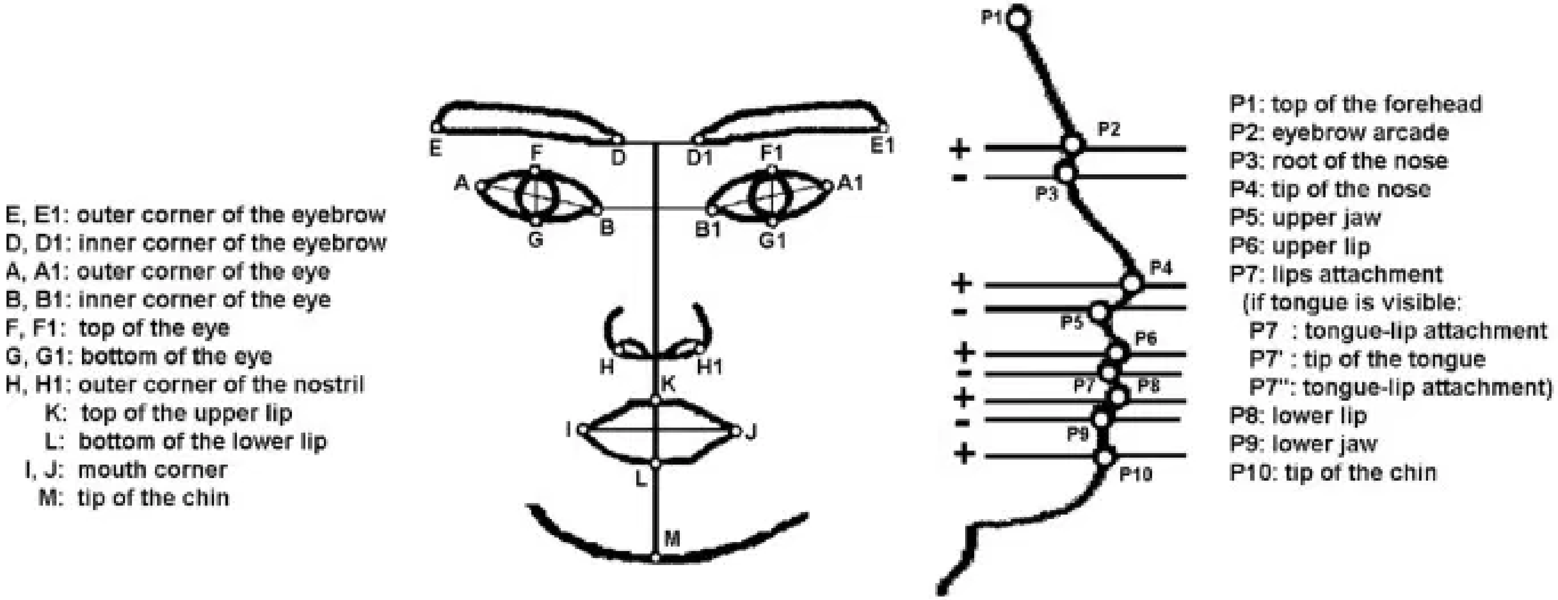
FEATURE REPRESENTATION FOR IMAGES IS COMPLEX



0	18	238	255	248	134	21	76	248	255	255	162	0	0	0	0	0	0	0	0	0	0	0	0
0	87	255	255	255	255	255	255	254	253	255	87	0	0	0	0	0	0	0	0	0	0	0	0
0	75	255	255	249	253	255	254	253	255	255	132	50	41	0	0	0	0	79	143	27	0	0	0
0	39	255	255	255	245	200	186	187	192	222	255	255	255	51	0	43	202	255	255	75	0	0	0
0	67	255	255	239	162	154	174	179	172	155	197	227	236	164	166	255	255	255	255	39	0	0	0
0	146	255	255	173	159	206	196	194	200	184	155	241	253	255	255	240	217	250	255	44	0	0	0
54	255	255	242	167	193	192	193	192	205	172	150	247	255	248	186	163	166	163	255	178	0	0	0
167	255	250	255	189	157	196	207	199	167	141	186	219	248	185	156	195	204	171	165	255	207	92	33
193	255	251	255	245	179	153	138	144	190	241	255	227	175	169	202	193	192	193	168	239	255	255	174
154	255	247	207	214	233	198	190	235	255	255	255	219	160	178	197	193	197	177	168	252	255	236	30
54	255	225	207	245	255	255	254	210	181	177	231	220	146	190	198	196	199	145	203	255	255	73	0
0	66	215	255	252	255	245	165	154	187	179	156	230	154	125	192	184	151	192	253	255	235	12	0
0	0	82	255	255	253	179	169	205	192	205	156	197	253	174	159	163	220	254	253	255	255	75	0
0	0	39	255	255	224	161	198	192	192	204	154	192	255	255	214	197	255	251	249	254	255	193	0
0	0	104	255	255	206	159	203	192	197	190	152	235	255	240	195	241	255	251	255	255	251	118	0
0	74	255	255	255	227	142	189	202	186	145	206	255	255	206	222	255	251	255	190	36	0	0	0
95	255	255	249	251	254	200	153	167	171	215	254	252	255	183	234	255	255	170	0	0	0	0	0
103	255	252	253	255	255	255	233	221	245	255	250	250	255	193	164	254	150	0	0	0	0	0	0
0	206	255	255	172	175	255	255	255	253	249	249	253	255	255	62	0	0	0	0	0	0	0	0
0	26	144	70	0	0	106	255	255	250	250	255	255	240	172	18	0	0	0	0	0	0	0	0
0	0	0	0	0	0	8	230	255	250	255	239	77	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	198	255	255	242	37	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	67	231	204	36	0	0	0	0	0	0	0	0	0	0	0	0	0

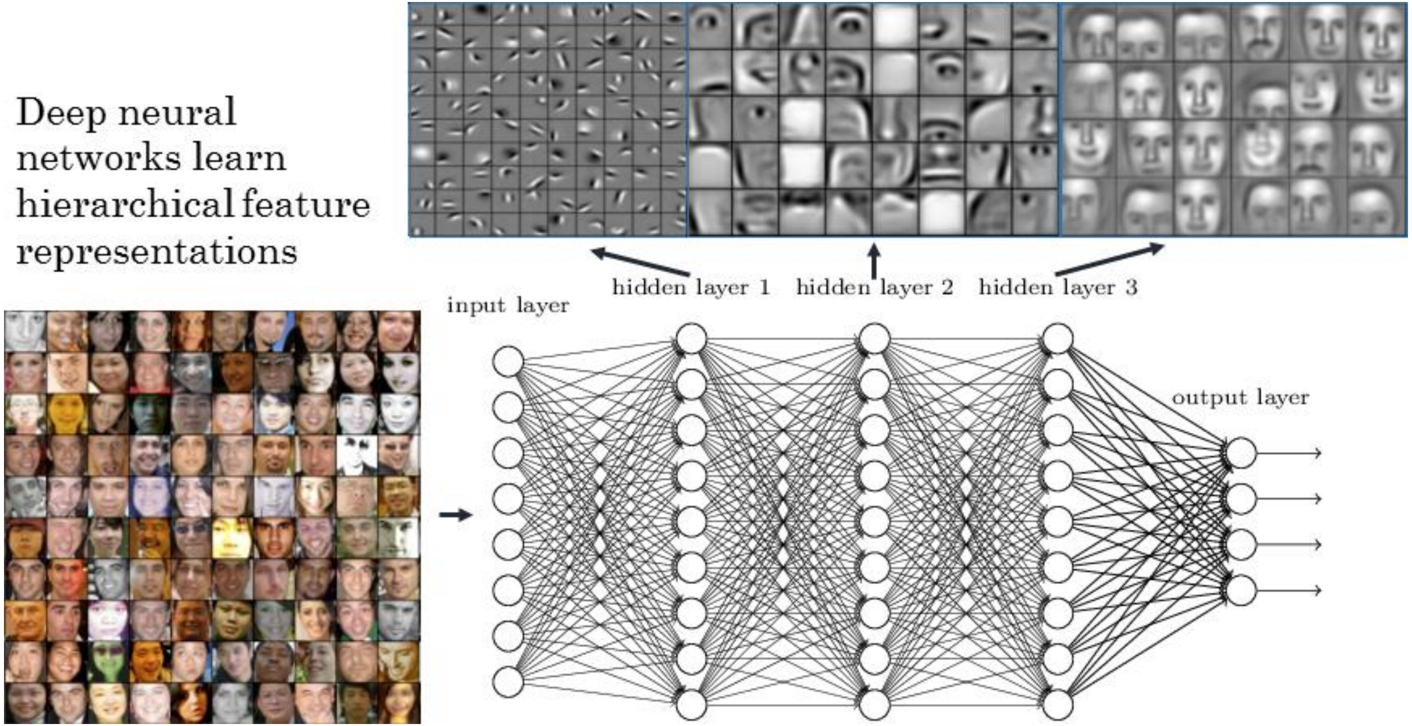
- Consider a face image of size 200 x 200 pixels
- A color (RGB) image has three channels
- Total pixels in the image are 200 x 200 x 3
- Feature extraction and representation from 120000 pixels per image becomes exceedingly complex
- Too many combinations to consider when trying to hand-craft features from so many pixels manually for machine learning

MANUAL FEATURE Extraction FOR IMAGES IS COMPLEX



DEEP LEARNING ENABLES AUTOMATED FEATURE EXTRACTION with ease

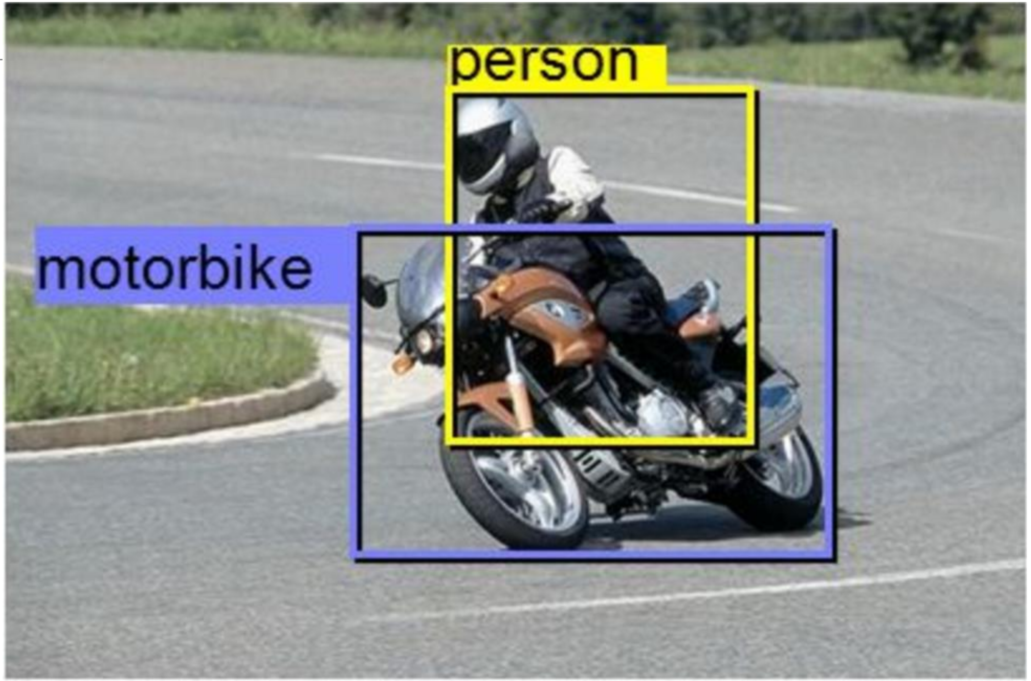
Deep neural networks learn hierarchical feature representations



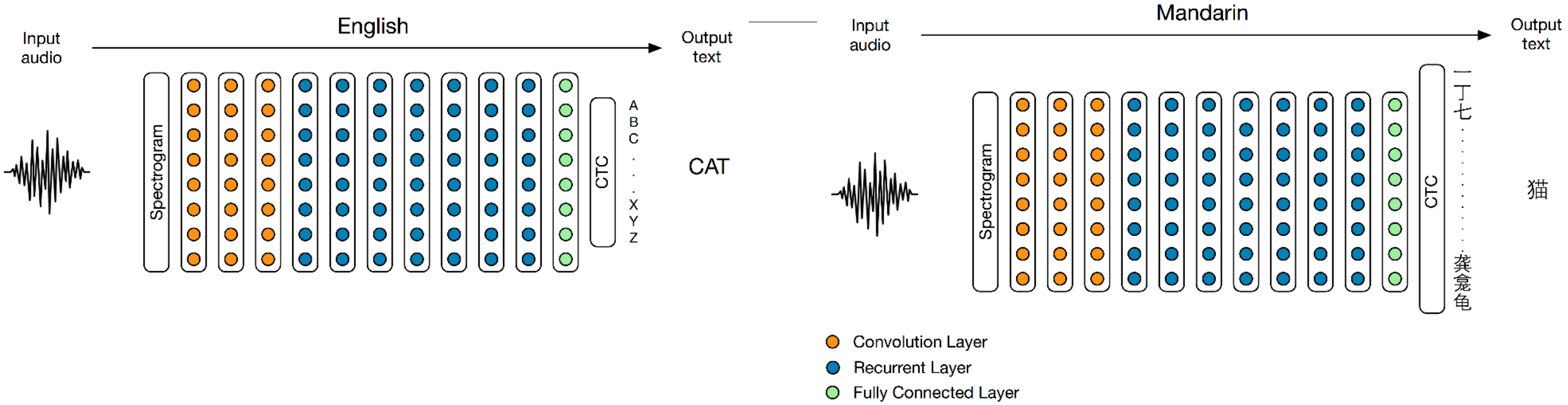
Classification and Detection

Detect and label the image

- Person
- Motor Bike



Speech Recognition and Language Translation



The same architecture is used for English and Mandarin Chinese speech recognition

http://svail.github.io/mandarin/

Scale drives deep learning progress