Introduction to Machine Learning

CSCS11H3F

Faiza Khan Khattak
University of Toronto Scarborough

Lecture Plan

- Course information and logistics
- Introduction to Machine Learning
  - Definitions and related terms
  - Types of ML

Teaching staff

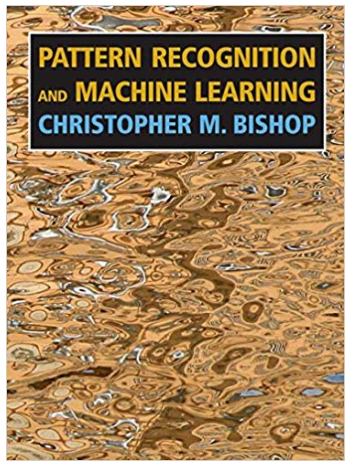
- **Faiza K. Khattak** (Instructor)
  Lecture: Wednesday (2PM-3PM), Friday (12PM-1PM)
  OH: Wednesday (3:00PM-4:00PM), Friday (1:30PM-2:30PM)
  Email: faiza.khattak@utoronto.ca
- **Ali Mojdeh** (TA)
  Tutorials: Tuesday (TUT0001) and (TUT0004),
  Email: ali.mojdeh@mail.utoronto.ca
- **Yuan Bian** (TA)
  Tutorial: Friday (TUT0002)
  OH: TBD
  Email: yuan.bian@mail.utoronto.ca

Prerequisites

- **Linear algebra:** vector, matrix manipulations, properties.
- **Calculus:** derivatives, partial derivatives/gradient.
- **Probability:** basic concept, common distributions, Bayes Rule.
- **Statistics:** expectation, variance, covariance, median, maximum likelihood.
- **Programming:** Python and libraries such as numpy and pandas.

Books

- **Textbooks:** **Pattern Recognition and Machine Learning** by Christopher Bishop, 2011.



- **Reference books:**
  - **Machine Learning: a Probabilistic Perspective** by Kevin Murphy, 2012.
  - **Information Theory, Inference, and Learning Algorithms** by David Mackay, 2003.

Other resources

- **Quercus:** Slides, HW and solutions will be posted to Quercus.
- Reference: Lecture notes by David Fleet (also posted on Quercus)
- **Piazza:** Used for discussion.
  Sign up link:
  `http://piazza.com/mail.utoronto.ca/fall2019/cscc11`
  Course link:
  `http://piazza.com/mail.utoronto.ca/fall2019/cscc11/home`

Assignments & Exams

- **Assignments:**
  Three worth 15% each, for a total of 45%. Will consist of theoretical and programming questions.
- **Mid-term:**
  In-class midterm in the week of Oct. 21. Worth 25% of course mark.
- **Final:**
  Worth 30% of course mark.

Rules

- Collaboration on the assignments is not allowed, unless otherwise stated. Each student is responsible for his/her own work.
- Discussion of assignments should be limited to clarification of the handout itself, and should not involve any sharing of answers to HW questions, pseudocode or code or simulation results.
- Violation of this policy is grounds for a semester grade of F, in accordance with university regulations.
- Assignments should be handed in by 11:59 pm on the due date. A late penalty of 10% per day will be assessed thereafter (up to 3 days, then submission is blocked).
- Extensions will be granted only in special situations (may need a Student Medical Certificate, if applicable).

Course Plan (Provisional)

- Overview of Machine Learning topics
- 1D regression, multidimensional regression, least-squares, pseudo-inverse
- Basis function regression, Radial Basis Functions, K-nearest nieghbours
- N-Fold Cross Valiadation, LOOCV
- Bayes' rule, Maximum likelihood, Maximum a Posteriori
- k-NN classifiers, Decision trees,
- Class conditional models,
- Logistic regression, Naïve Bayes
- SVM Kernels

Course Plan (Provisional)

- Entropy, Mutual Information, KL Divergence, Cross-Entropy
- Ensemble learning (Boosting and Bagging)
- Bayesian Regression, Model Averaging, Model Selection
- Sampling Gaussians and Categorical Distribution, Importance Sampling, MCMC
- Dimensionality Reduction, PCA, Probabilistic PCA
- K-means, Gaussian Mixture Models,
- Expectation-Maximization Algorithm
- Perceptron
- Deep Learning

What is Machine Learning?

Machine Learning and AI in movies

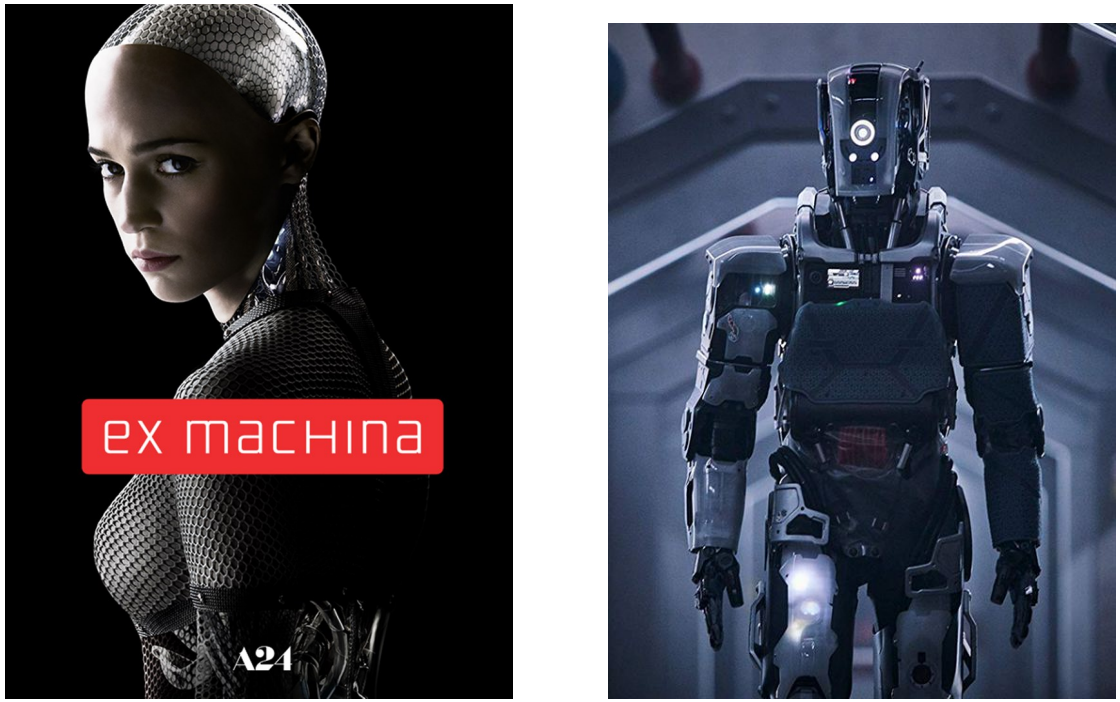


Figure: (Left): Image from movie "Ex Machina" (Right): Image from movie "I Am Mother"

Building Intelligent Machines

- For years scientists have been trying to invent machines that can automate the tasks, especially the tasks that are tedious, difficult, and time consuming.
- Make the machines solve problems just like or better than the way humans do e.g., classifying images, predicting the stock market, and obey orders from human.

Building Intelligent Machines

- For years scientists have been trying to invent machines that can automate the tasks, especially the tasks that are tedious, difficult, and time consuming.
- Make the machines solve problems just like or better than the way humans do e.g., classifying images, predicting the stock market, and obey orders from human.

Common ways of creating intelligent machines include:

- Rule based methods
- Machine learning methods

Building Intelligent Machines

Rule based methods

Perform a task based on defined rules.

Toy example: Classify a table vs. not a table.

Input: If an object:

- **Rule 1:** is made of wood, glass or plastic.
- **Rule 2:** has four legs.
- **Rule 3:** has a square, rectangular or round shape .

$$\Longrightarrow \text{ It is a } \textbf{table}$$

Building Intelligent Machines

Rule based methods

Perform a task based on defined rules.

Toy example: Classify a table vs. not a table.

Input: If an object:

- **Rule 1:** is made of wood, glass or plastic.
- **Rule 2:** has four legs.
- **Rule 3:** has a square, rectangular or round shape .

$\Longrightarrow$ It is a **table**

Output:

Table

Building Intelligent Machines

Rule based methods

Perform a task based on defined rules.

Toy example: Classify a table vs. not a table.

Input: If an object:

- **Rule 1:** is made of wood, glass or plastic.
- **Rule 2:** has four legs.
- **Rule 3:** has a square, rectangular or round shape .

$\Longrightarrow$ It is a **table**

Output:

Table

Building Intelligent Machines

Rule based methods

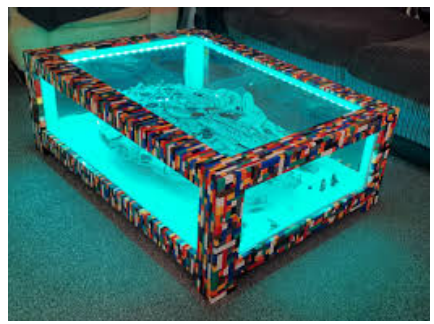
- Rules have to be well-defined, comprehensive, and exhaustive.
- Encoding everything in the form of rules can be difficult, especially for complicated concepts.

Building Intelligent Machines

Machine learning methods:

- Instead of explicitly writing all rules, use examples to train the computer to perform this task.
- Show the computer thousands/millions of examples of tables.
- Computer learns automatically to classify a table.






Some definitions

Artificial Intelligence (AI)

AI is the study of building systems that can imitate intelligent human behavior. This can include rule-based methods and machine learning methods.

Some definitions

Artificial Intelligence (AI)

AI is the study of building systems that can imitate intelligent human behavior. This can include rule-based methods and machine learning methods.

Machine Learning (ML)

Machine learning consists of development and use of algorithms that learn and improve automatically through experience based on data.

Some definitions

Artificial Intelligence (AI)
AI is the study of building systems that can imitate intelligent human behavior. This can include rule-based methods and machine learning methods.

Machine Learning (ML)
Machine learning consists of development and use of algorithms that learn and improve automatically through experience based on data.

Deep Learning (DL)
A subset of ML algorithms where artificial neural networks, which are algorithms inspired by the human brain, learn from large amounts of data.

Some definitions

Artificial Intelligence (AI)
AI is the study of building systems that can imitate intelligent human behavior. This can include rule-based methods and machine learning methods.

Machine Learning (ML)
Machine learning consists of development and use of algorithms that learn and improve automatically through experience based on data.

Deep Learning (DL)
A subset of ML algorithms where artificial neural networks, which are algorithms inspired by the human brain, learn from large amounts of data.

Data Mining (DM)
Data mining refers to knowledge extraction from large amounts of data. This includes discovering various types of useful patterns and information in data using ML algorithms.

Graphical view

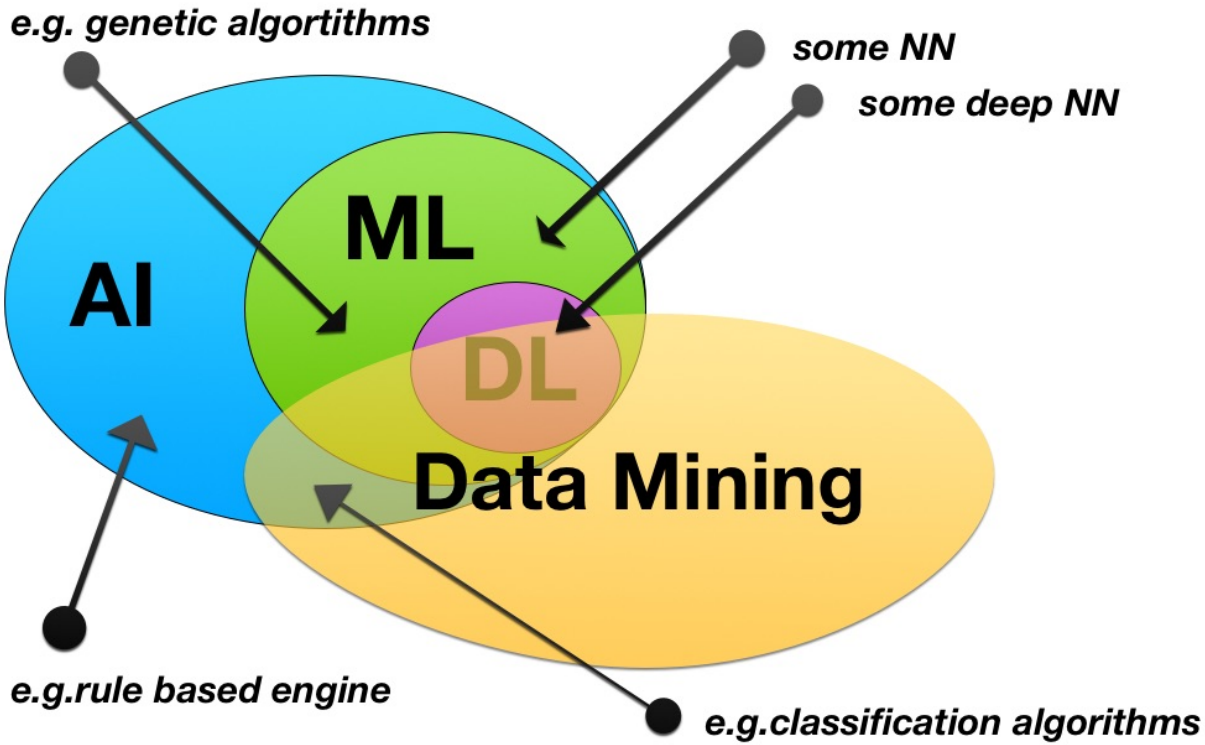
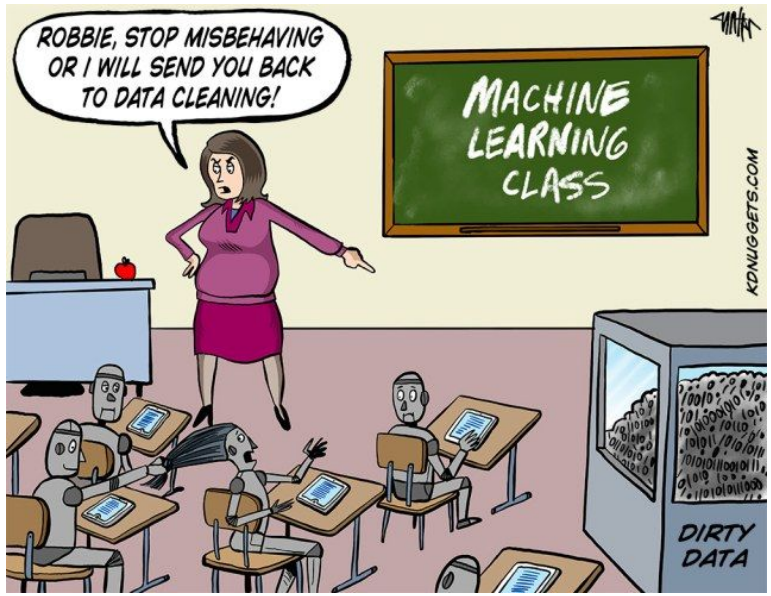


Image from `https://softwareengineering.stackexchange.com`

The Data

- ML algorithms learn from data.
- Data has useful information but most of the data is **messy**, **noisy**, and **unlabeled**.
- Data has to be cleaned and converted to a format that can be used by any machine learning algorithms.
- Data cleaning can be a tedious task and may take a lot of time.



Datasets, Features and Labels

- **Data instances:** Datasets consists of data instances usually denoted by $\{\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_n}\}$ e.g., houses for sale.

Datasets, Features and Labels

- **Data instances:** Datasets consists of data instances usually denoted by $\{\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_n}\}$ e.g., houses for sale.
- **Labels:** Labels are the tags or information provided about the data instances, usually denoted by $\{y_1, y_2, \cdots, y_n\}$ e.g., prices of houses for sale. However, a machine learning algorithm may not always depend on labels.

Datasets, Features and Labels

- **Data instances:** Datasets consists of data instances usually denoted by $\{\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_n}\}$ e.g., houses for sale.
- **Labels:** Labels are the tags or information provided about the data instances, usually denoted by $\{y_1, y_2, \cdots, y_n\}$ e.g., prices of houses for sale. However, a machine learning algorithm may not always depend on labels.
- **Features/attributes:** Features are the variables associated with data instances that describe the data instance and can help in building an accurate ML model e.g., to predict the price of a house the features could include size of the house, location of the house etc. The number of features is called **dimension**.

  ith instance $= \mathbf{x_i} = (x_i^{(1)}, x_i^{(2)}, \cdots, x_i^{(d)})$, where d is the dimension.

How do Machines Learn?

Data are randomly divided in **training set** and **test set**, usually 80% and 20% of the data respectively.

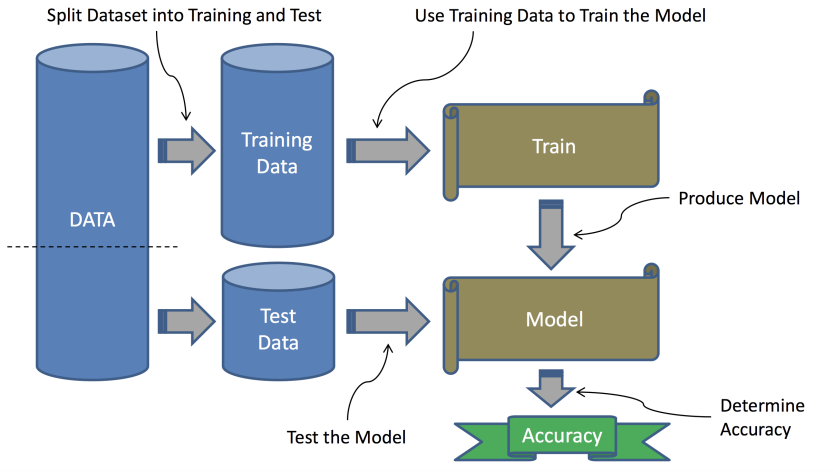


Image from `https://www.slideshare.net/AndrewFerlitsch/`

How do Machines Learn?

- **Training phase:** In this phase, training set examples are input in the ML algorithm to learn.
- **Testing phase:** In this phase, test set examples (unseen examples) are used to analyze how well the model has learned during the training phase.
- **Model:** A machine learning model is a mathematical representation of what was learned from training data.

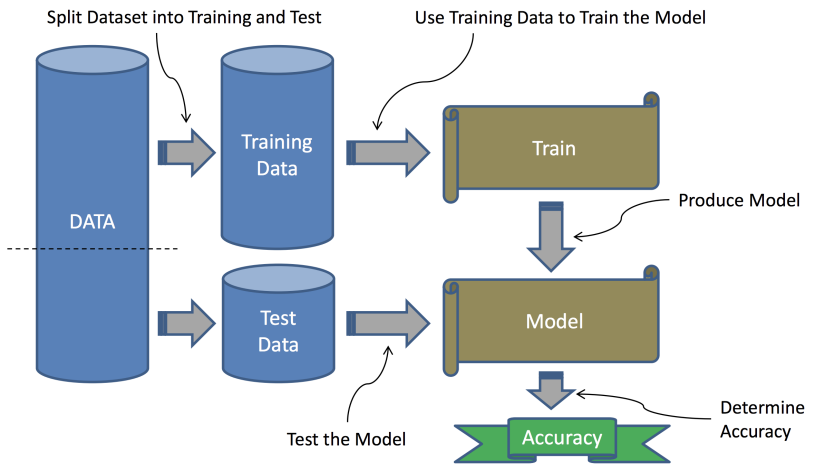


Image from https://www.slideshare.net/AndrewFerlitsch/

Main categories of Machine Learning

Main categories of Machine learning are:

1. Supervised Learning
2. Unsupervised learning
3. Reinforcement learning

Main categories of Machine Learning

1- Supervised Learning

- Supervised learning happens when the data has labels (sometimes called true labels or ground truth), which are used for learning.
- During the training phase the algorithm gets the labeled data as input to learn the model.
- In the testing phase model is tested on the data without labels provided.
- Accuracy is checked by comparing the true labels and output labels.
- **Examples:**
  - **Classification:** The outputs are discrete labels, as in email spam filtering.
  - **Regression:** The outputs are real-valued e.g., predicting prices of the houses.

Main categories of Machine Learning

1- Supervised Learning

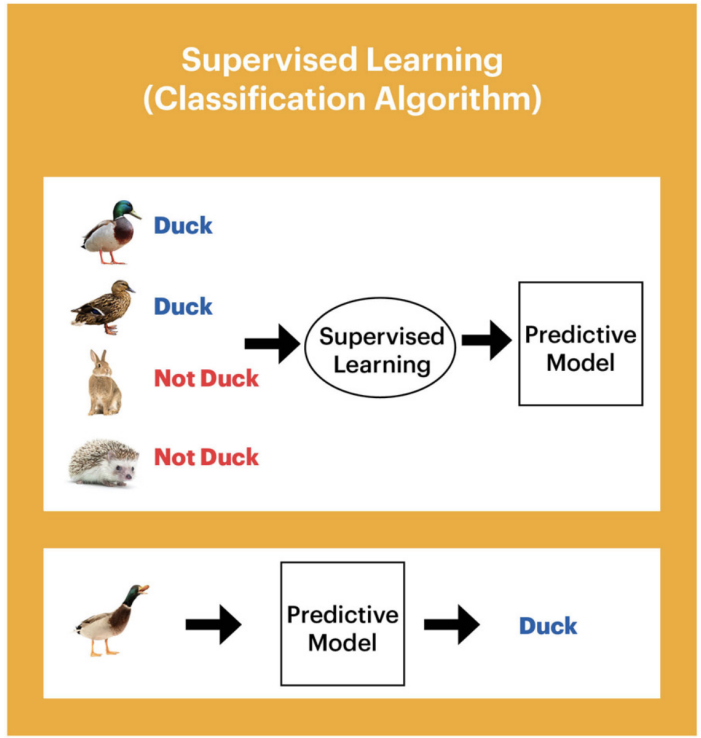


Image from https://blog.westerndigital.com/machine-learning-pipeline-object-storage/

Main categories of Machine Learning

2- Unsupervised Learning

- A collection of unlabeled data is used to analyze and discover patterns within.
- **Example:**
  - **Clustering:** Dividing a set of data points into clusters, where data points assigned to the same cluster are similar, and data points assigned to different clusters are dissimilar.

Main categories of Machine Learning

2- Unsupervised Learning

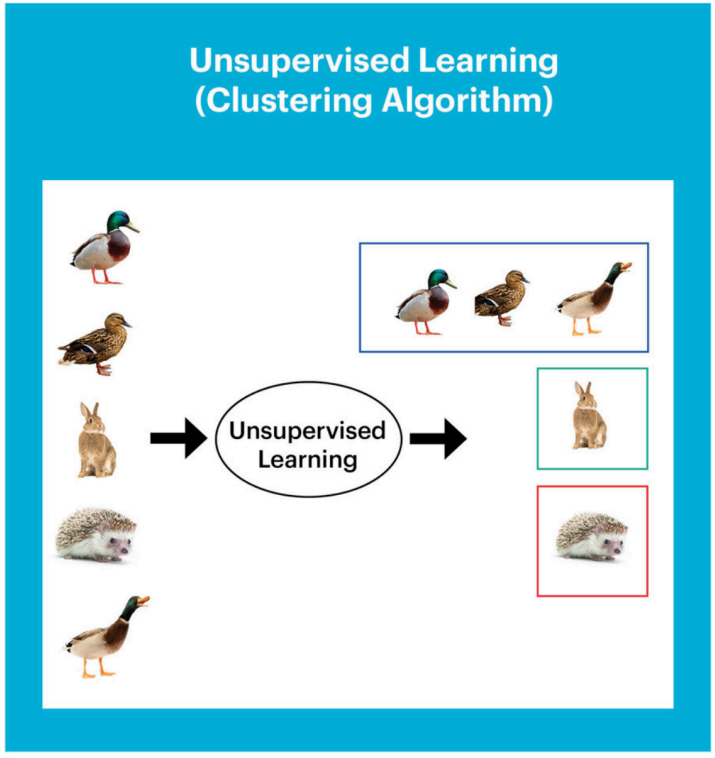


Image from https://blog.westerndigital.com/machine-learning-pipeline-object-storage/

Main categories of Machine Learning

3- Reinforcement Learning

An agent (e.g., a robot or controller) seeks to learn the optimal actions to take, based the outcomes of past actions. Outcomes are provided in sequences based on decisions.
There are five elements associated with reinforcement learning:

- **Agent:** an intelligent program and decision-maker.
- **Environment:** the surrounding area, which has a goal for the agent to perform.
- **Internal state:** maintained by an agent to learn the environment.
- **Actions:** the tasks carried out by the agent in an environment.
- **Rewards:** used to train the agents.

Example: Chess game.

Main categories of Machine Learning

3- Reinforcement Learning

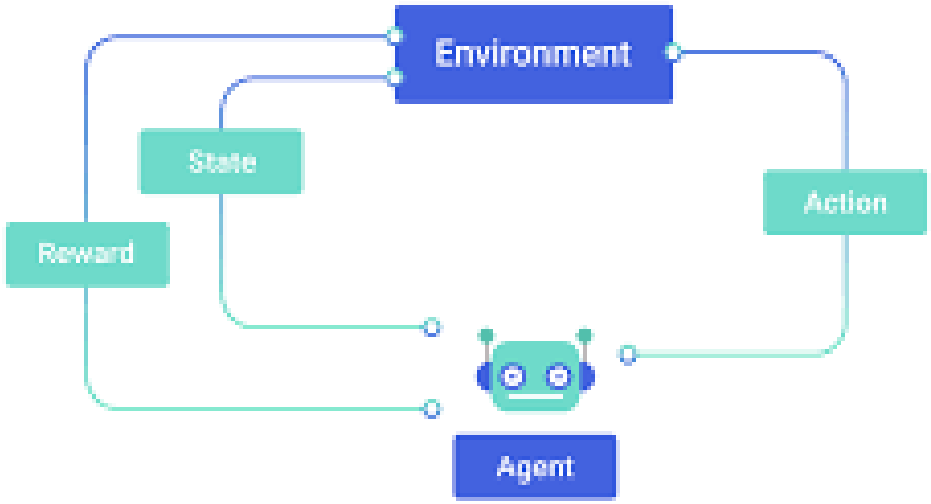


Image from https://becominghuman.ai/the-very-basics-of-reinforcement-learning-154f28a79071

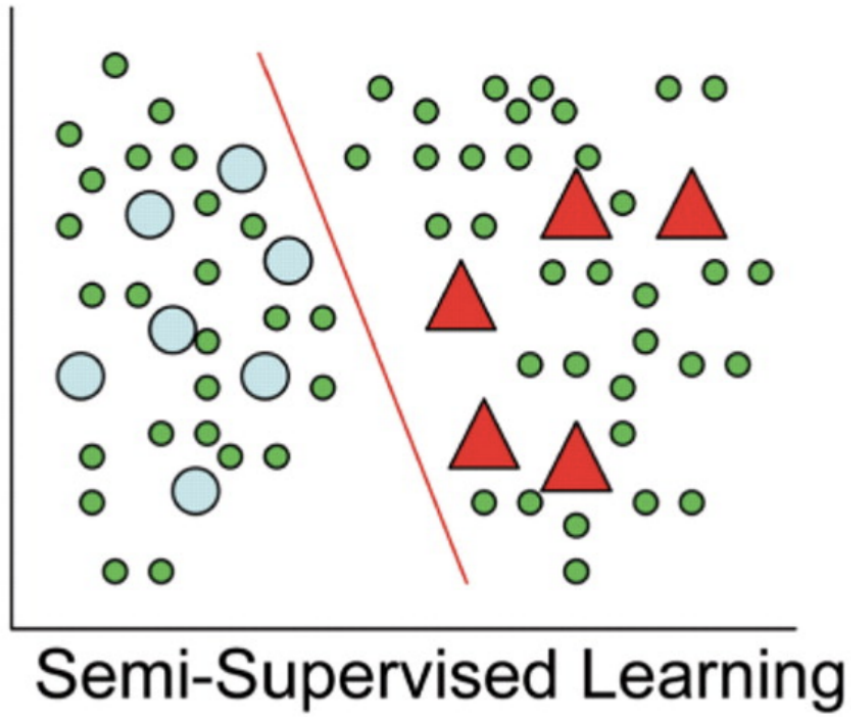
Other types of Machine Learning

Other types of ML include:

- Semi-supervised learning
- Active learning
- Transfer Learning
- Structured Prediction
- Time-series forecasting
- Anomaly detection
- Deep Learning

Other types of Machine Learning

Semi-supervised Learning: Only a subset of the training data is labeled.



Other types of Machine Learning

Active Learning: Obtaining labeled data is expensive (usually has to be acquired by human experts.) An algorithm can determine which training data to acquire.

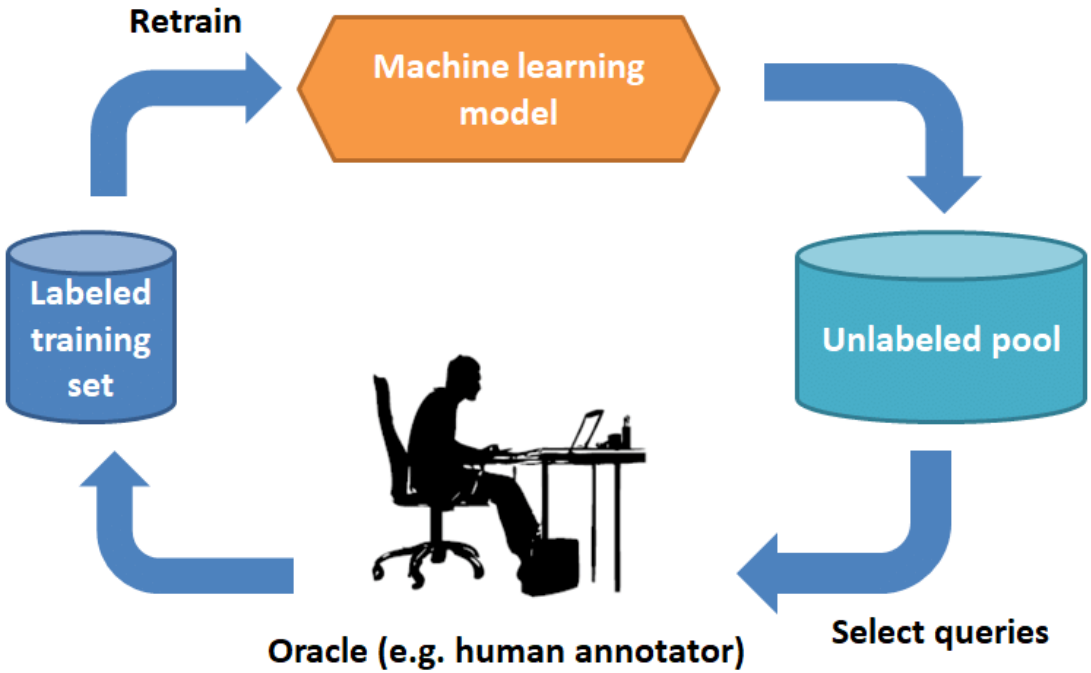


Image from https://www.researchgate.net/figure

Other types of Machine Learning

Transfer Learning: Models learned from tasks with large training sets can be used to help constrain learning on related problems that have relatively small datasets.

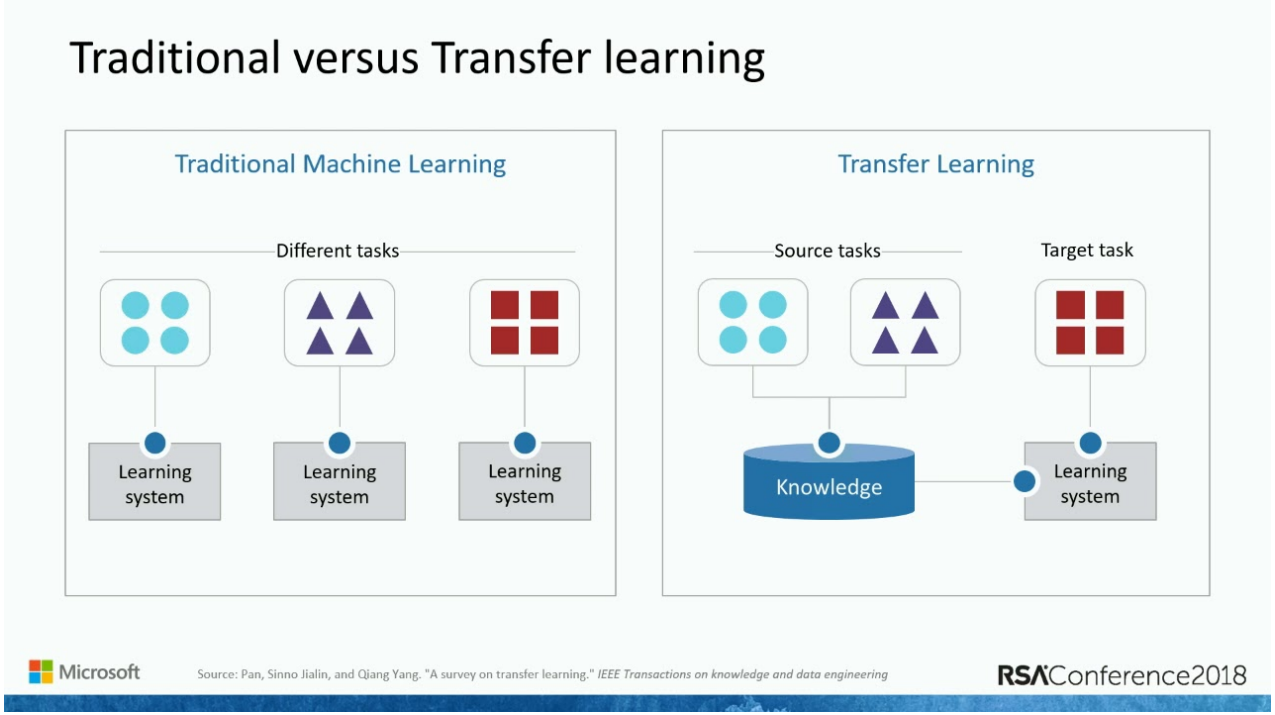
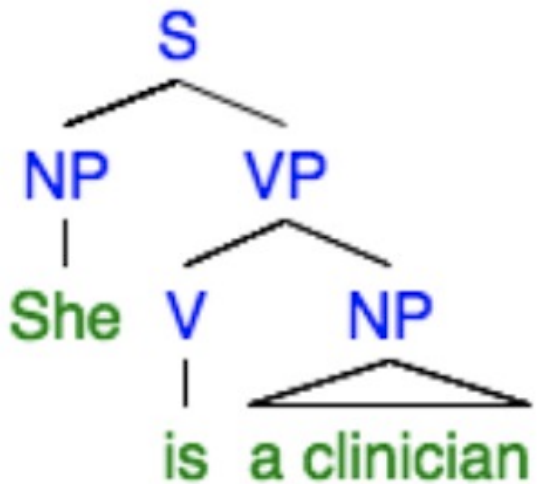


Image from slides by Mark Russinovich, Chief Technology Officer, Azure, Microsoft

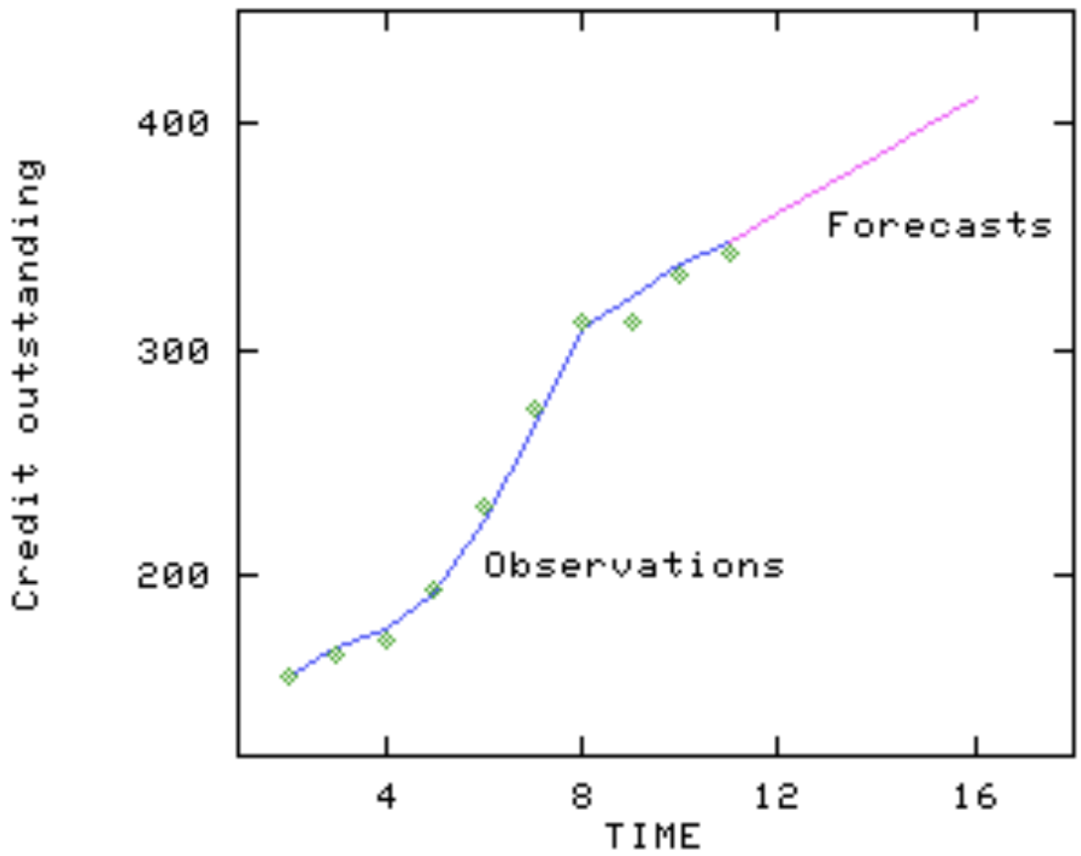
Other types of Machine Learning

Structured Prediction: Learn models for complex structured objects, rather than scalar discrete or real values. The structured objects include graphs or images, with many dimensions and complex dependencies among the predicted variables.
Example: Translating a natural language sentence into a syntactic representation such as a parse tree.
Text: 'She is a clinician'
S: Sentence, **NP:** Noun Phrase, **VP:** Verb Phrase, **V:** Verb



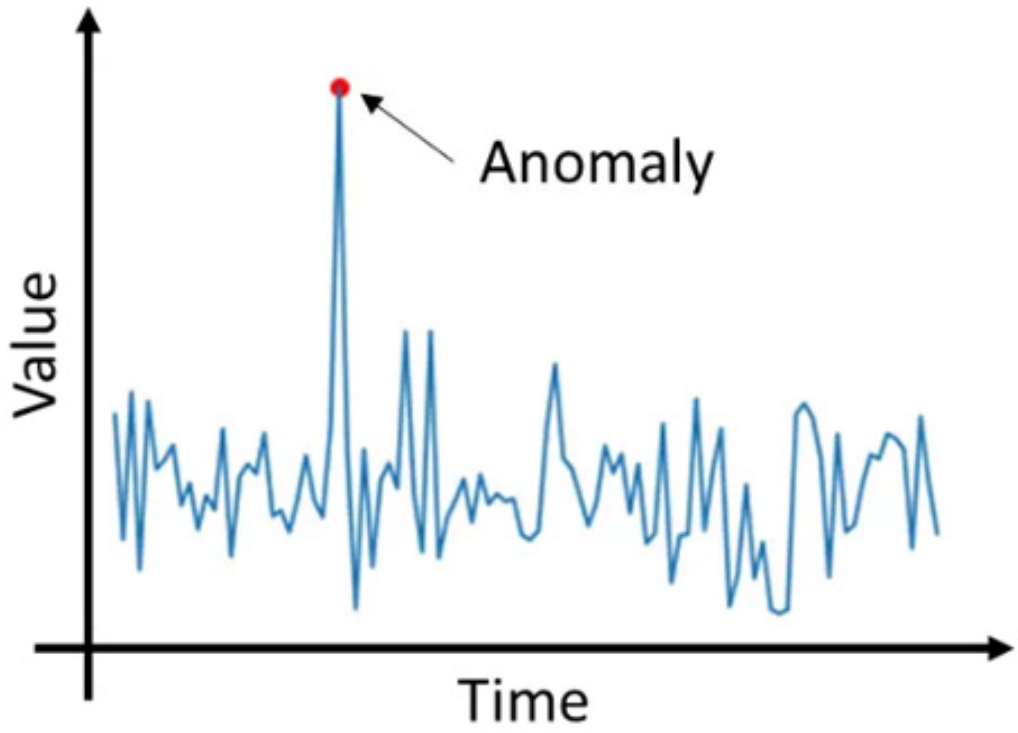
Other types of Machine Learning

Time-series forecasting: Predicting future trends such as in financial markets.



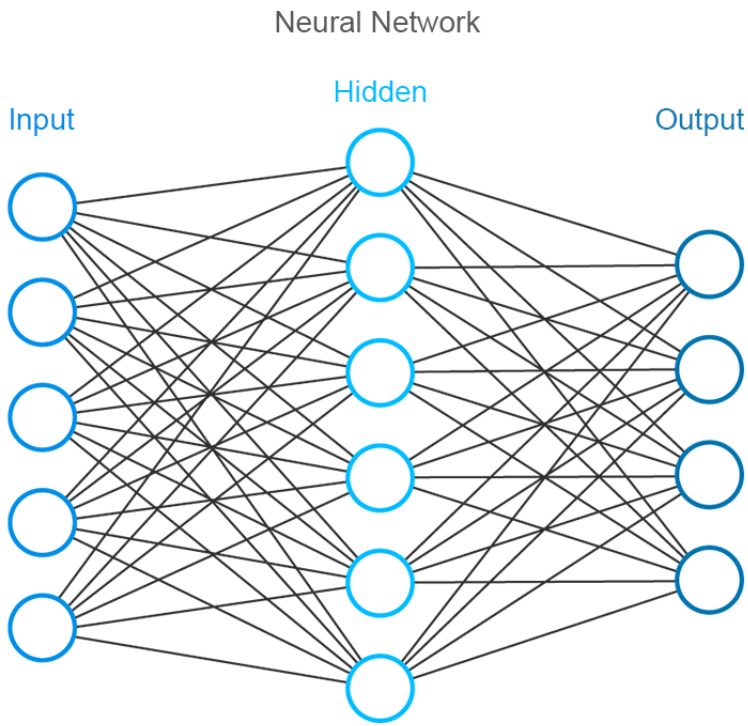
Other types of Machine Learning

Anomaly detection: Identifying the rare observations such as fault-detection in factories or fraud detection in banks.



Other types of Machine Learning

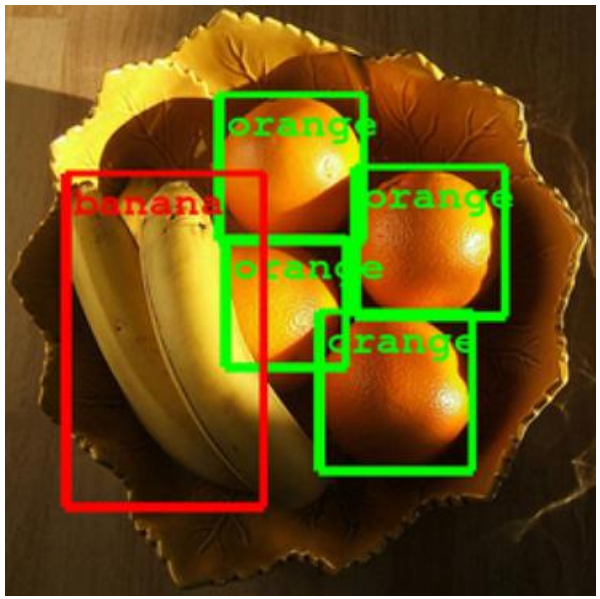
Deep Learning: Inspired by human brain neuron. Concerns the design and application of multilayer artificial neural networks. Had great success in a wide range of supervised learning tasks.



Applications of Machine Learning

Computer vision: Make computer recognize or identify parts of the image or video. It includes:

- object detection i.e., detecting instances of a given class,
- semantic segmentation i.e.,inferring labels for every pixel in images,
- human pose estimation i.e., orientation of a person in an image,
- and a lot more...



Applications of Machine Learning

Speech recognition: Recognizing and converting of spoken language into text by computers. This includes:

- Speech to text for note taking,
- personal assistance responding to questions,
- speaker identification.

Example: Amazon's Alexa, Apple's Siri and Microsoft's Cortana.

Applications of Machine Learning

Natural Language Processing (NLP):

- Machine translation e.g., Google translate,
- sentiment analysis,
- topic modeling i.e., identifying underlying document topics,
- spam filtering.

Figure: Topic modeling of 1.8M New York times articles.

References

1. Lecture notes by Prof. David Fleet `http://www.cs.toronto.edu/~fleet/courses/C11/Handouts/Introduction.pdf`

2. "Pattern Recognition and Machine Learning" by Christopher Bishop