Introduction to
Machine Learning

Sami Ghawi & Mustafa Jarrar

Birzeit University

Watch this lecture and download the slides

Course Page: http://www.jarrar.info/courses/AI/

More Online Courses at: http://www.jarrar.info

Outline

Overview about Machine Learning and its paradigms and applications

- Introduction and Motivation
- Challenges of Machine Learning
- Learning Types
  - Supervised Learning
  - Unsupervised Learning
  - Reinforcement Learning
- Real World Examples

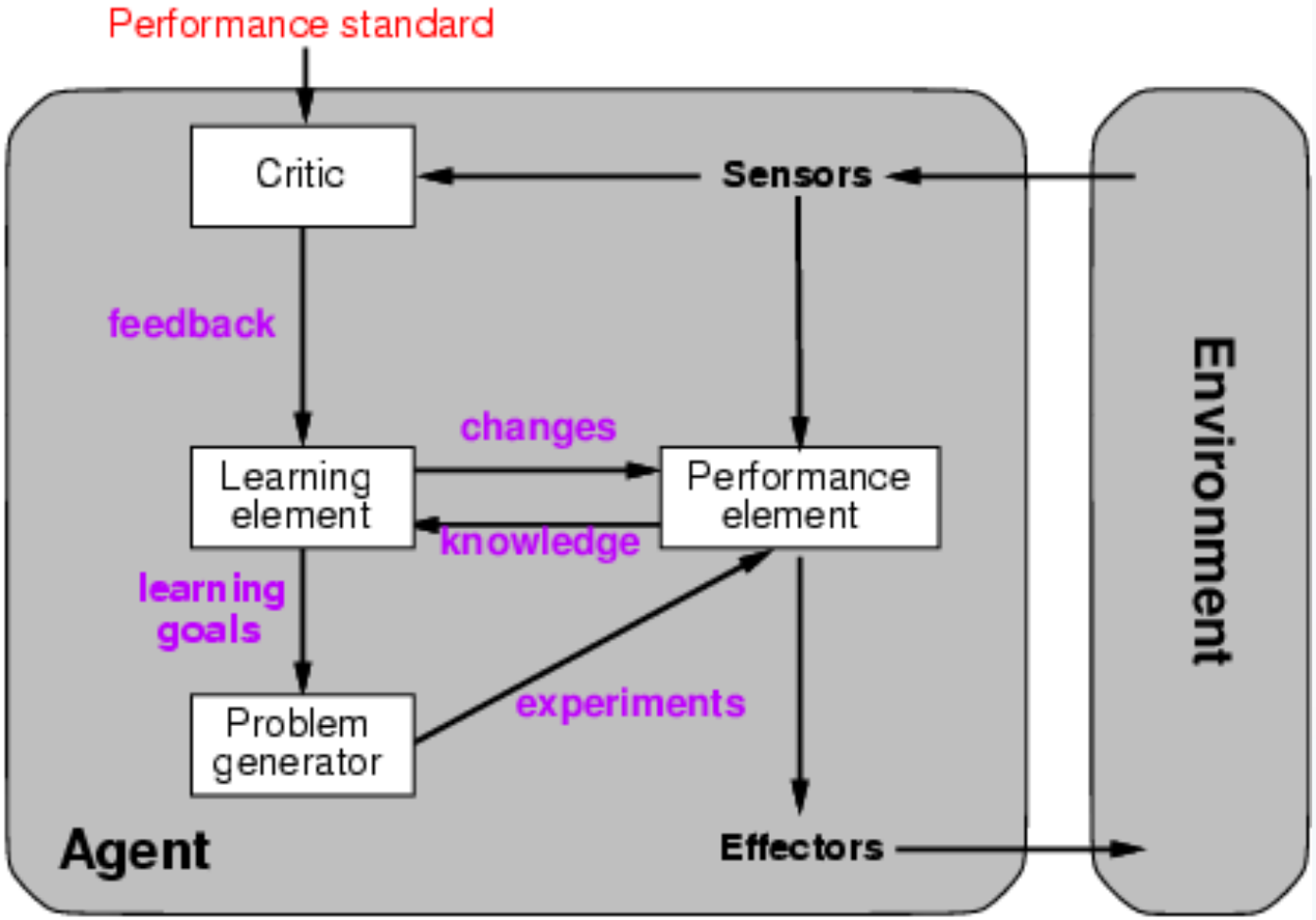
Keywords: Learning, Machine learning, Supervised Learning, unsupervised Learning, Reinforcement learning

التعلم الآلي، تعليم الآلة، التعليم الموجة، التعليم الغير موجه، التعليم المعزز

Learning Agents

Based on [8]

The agent adapts its action(s) based on feedback (not only sensors).



Introduction

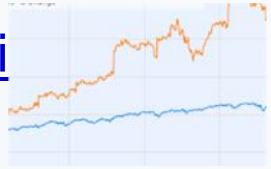
What is Machine Learning?

Field of study that gives computers the ability to learn without being explicitly programed (Arthur Samuel 1959)

Why is Machine Learning needed?

Machine Learning is used when [1,2]:

- Human expertise does not exist. (Curiosity Rover).
- Humans are incapable of explaining their expertise(Speech Recognition).
- Amount of data is too large for a human to analyze (Data Mini



- Prediction of new data (Stock Market Prediction).
- Tasks that are learnt by practicing (Robot Path Planning).

Motivation: Inductive Learning

Based on [8]

Simplest form: learn a function from examples

f is the target function

An example is a pair (*x*, *f(x)*)

Problem: find a hypothesis *h*
such that $h \approx f$
given a training set of examples

This is a highly simplified model of real learning:
- Ignores prior knowledge
- Assumes examples are given

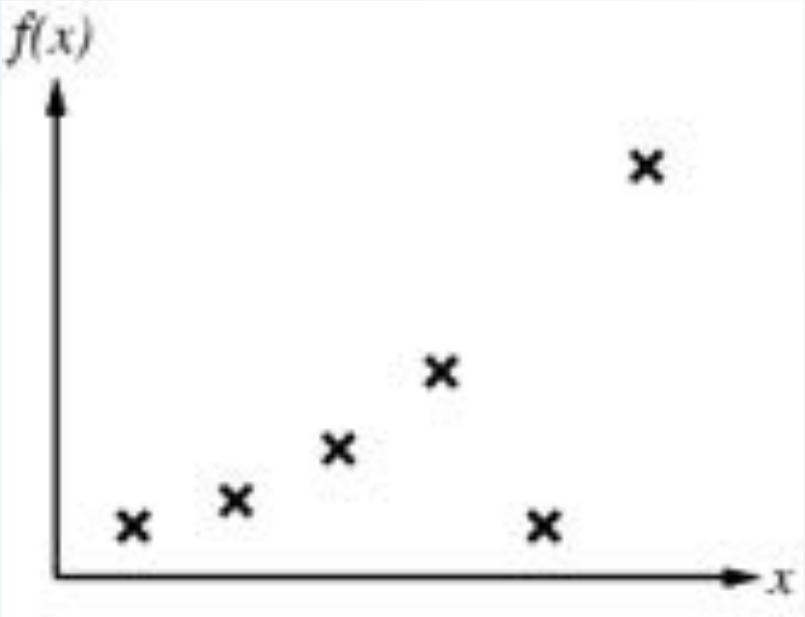
Motivation: Inductive Learning

Based on [8]

Construct/adjust *h* to agree with *f* on training set

(*h* is consistent if it agrees with *f* on all examples)

E.g., curve fitting:



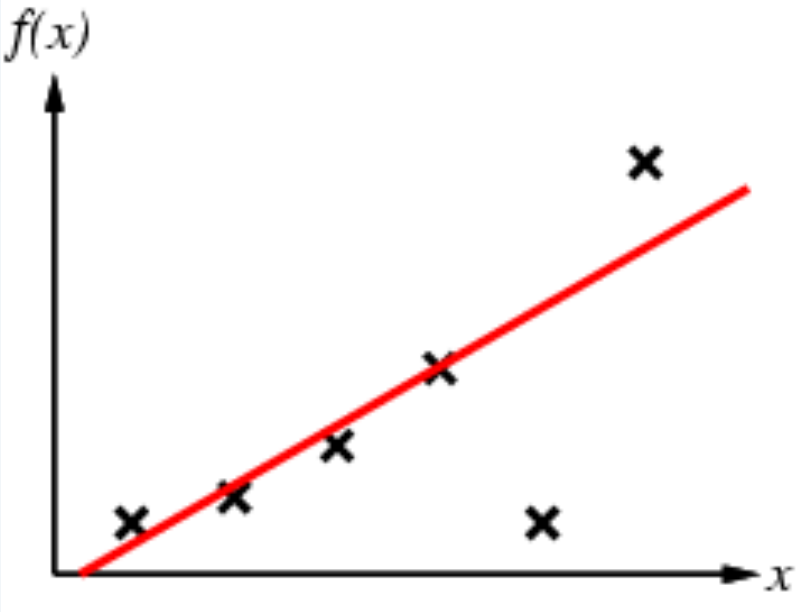
Motivation: Inductive Learning

Based on [8]

Construct/adjust *h* to agree with *f* on training set

(*h* is consistent if it agrees with *f* on all examples)

E.g., curve fitting:



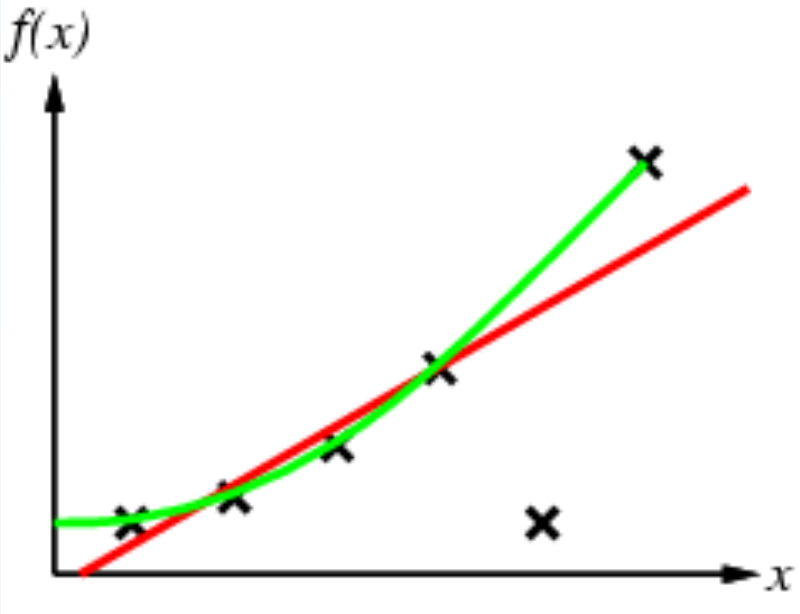
Motivation: Inductive Learning

Based on [8]

Construct/adjust *h* to agree with *f* on training set

(*h* is consistent if it agrees with *f* on all examples)

E.g., curve fitting:



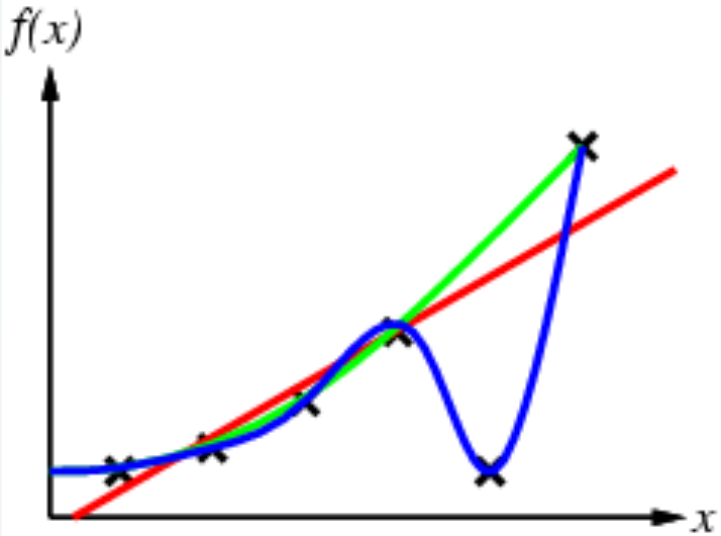
Motivation: Inductive Learning

Based on [8]

Construct/adjust *h* to agree with *f* on training set

(*h* is consistent if it agrees with *f* on all examples)

E.g., curve fitting:



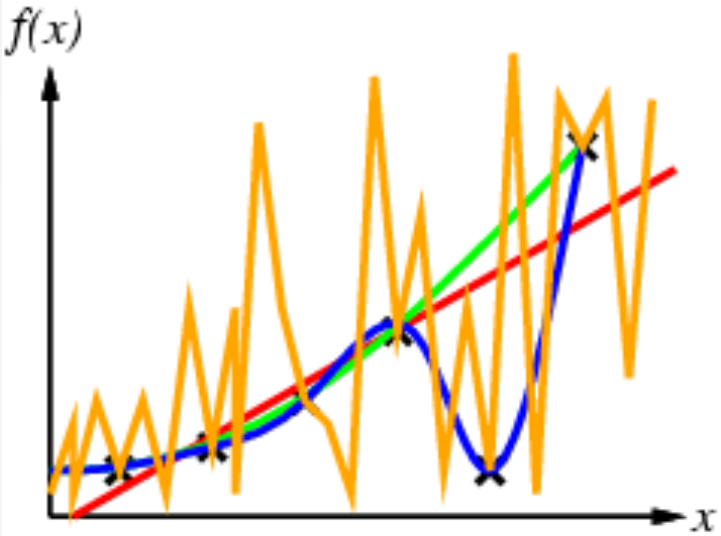
Motivation: Inductive Learning

Based on [8]

Construct/adjust *h* to agree with *f* on training set

(*h* is consistent if it agrees with *f* on all examples)

E.g., curve fitting:

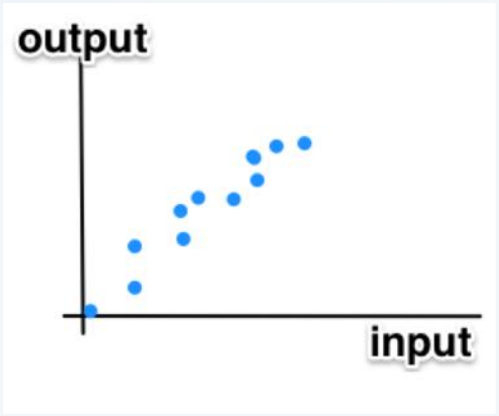


Ockham's razor: prefer the simplest hypothesis consistent with data

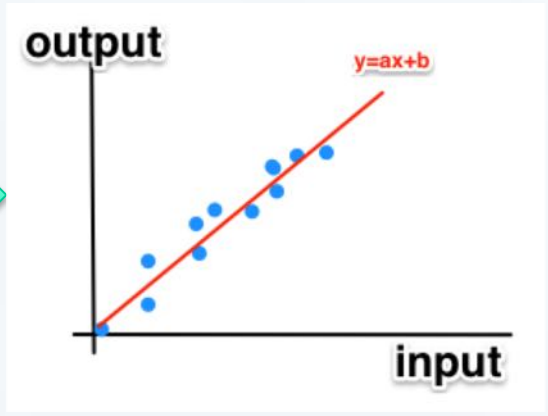
Introduction to Machine Learning

What is meant by learning?

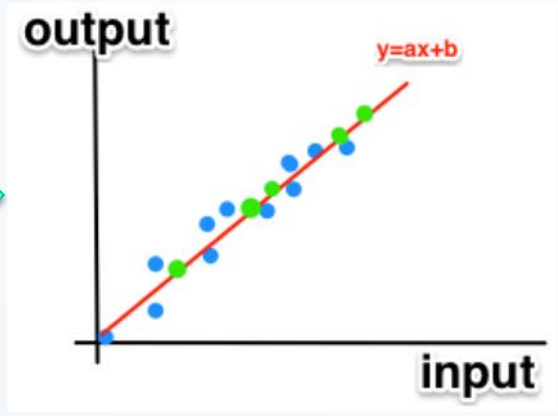
- Writing algorithms that can learn patterns from data.
- The algorithms create a statistical model that is a good approximation of the data.



Data from Past Experiences



Calculating a model



Estimating the output for new input values

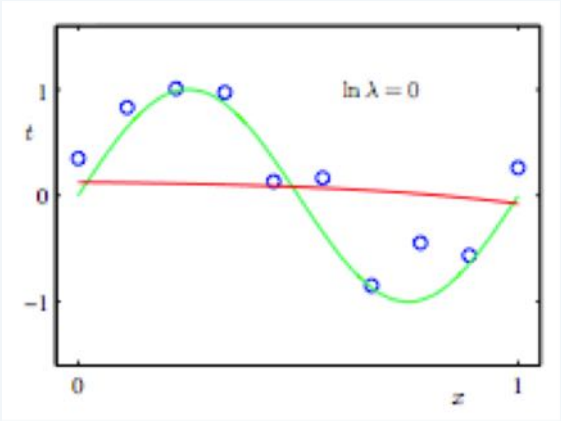
Challenges of Machine Learning

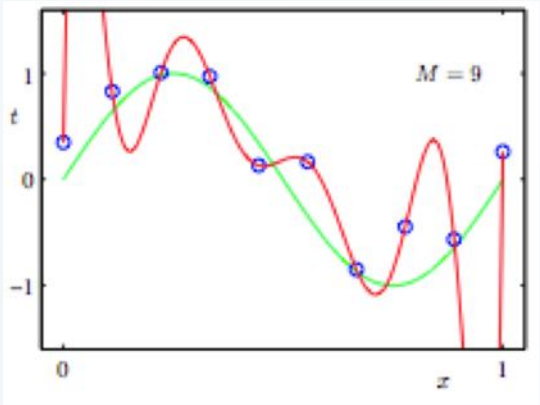
High Dimensionality [3]

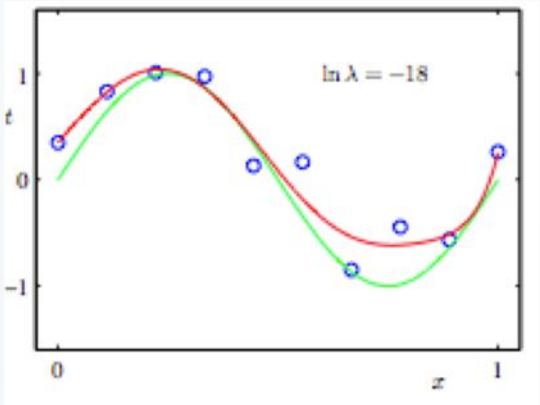
- Complexity of the data becomes very high and requires bigger models
- Requires a greater amount of memory and more time to process.
- Might cause over-fitting.
- Example: DNA Microarray

Choice of Statistical Model [4]

- Choosing the correct model and parameters that satisfy the available data
- Can cause under-fitting or over-fitting



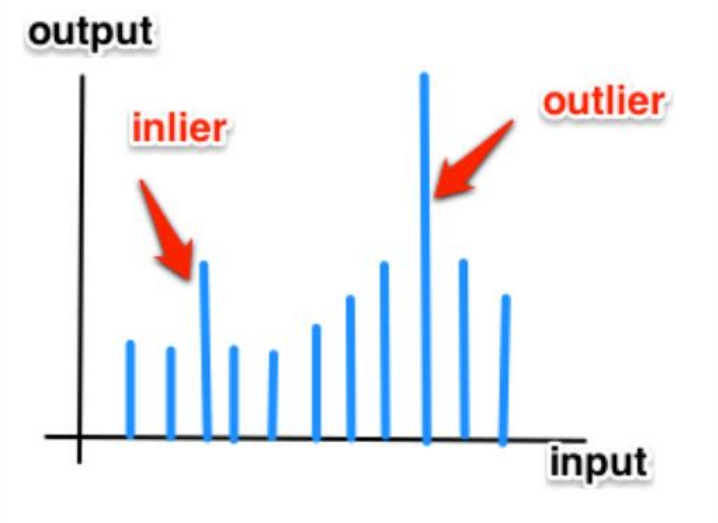




Challenges of Machine Learning

Noise and Errors [5]

- Gaussian Noise: Statistical Noise that has its probability density function equal to normal distribution.
- Outlier: an observation that is distant from the rest of the data.
- Inlier: a local outlier. (see: 2-sigma rule).
- Human Error causing incorrect measurements



Challenges of Machine Learning

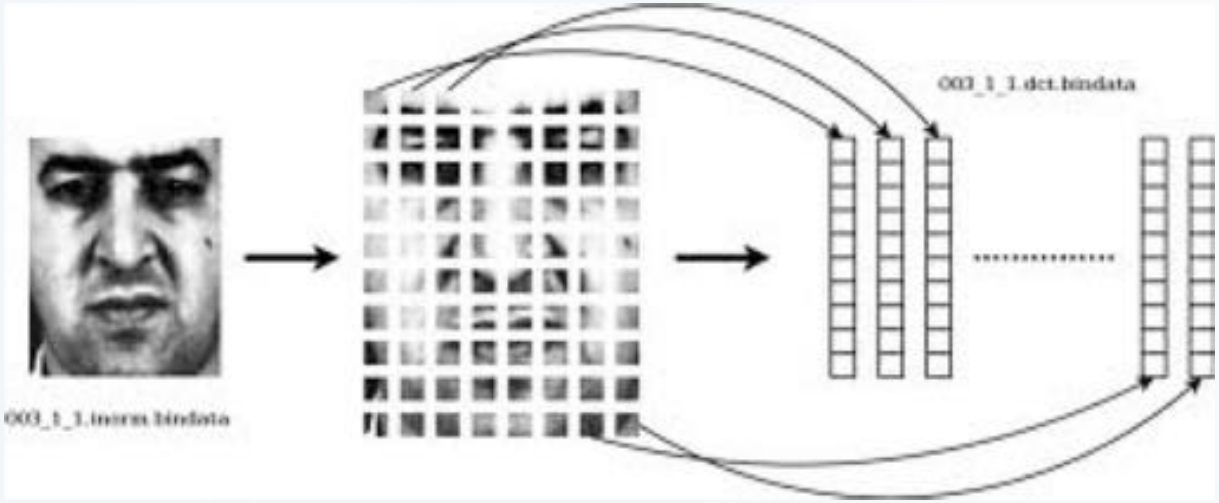
Insufficient Training Data

The amount of data is not sufficient to build a good approximation of the process that generated the data.

Feature Extraction in Patterns

Feature extraction is the process of converting the data to a reduced representation of a set of features.

Image Reference:
Face Verification



Learning Types

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

- Other Learning Paradigms

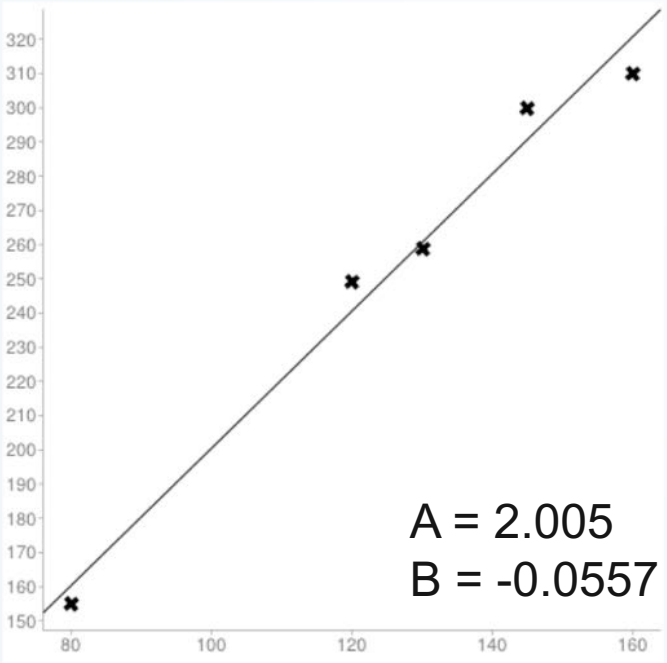
Supervised Learning

Regression:

- Regression aims to estimate a response.
- The output **y** takes numeric values.
- Toy Problem: We have a data of apartments with their areas and prices. We want to find a model that describe it and predict the prices of other areas (Assuming that all other variables don't have any effect).

Example of Training Data:

Area (m^2)	Price (1000$)
80	155
120	249
130	259
145	300
160	310



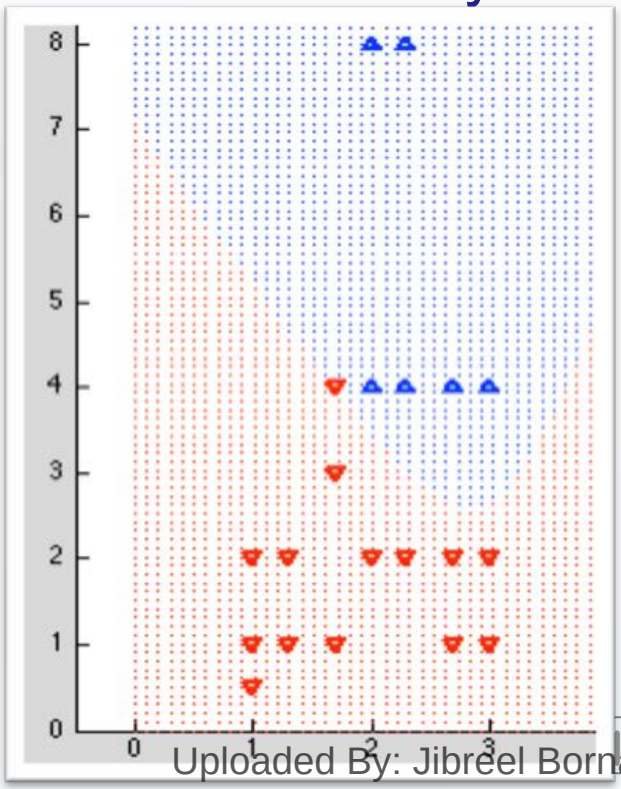
Supervised Learning

Classification:

- Classification aims to identify group membership.
- The output **y** takes class labels.
- Toy Problem: We want to determine whether a Computer is good or not from the processor and available memory

Example of Training Data:

Processor (GHz)	Memory (GB)	Status
1.0	1.0	Bad
2.3	4.0	Good
2.6	4.0	Good
3.0	8.0	Good
2.0	4.0	Bad
2.6	0.5	Bad



Unsupervised Learning

- Training data contain only the input vectors [4].
- Definition of training data: $\{x_1, x_2, ..., x_n\} \quad x \in R^A$
- Goal: Learn some structures in the inputs.
- Can be divided to two categories: Clustering and Dimensionality Reduction

Unsupervised Learning

Clustering

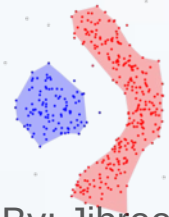
- Clustering aims to group input based on the similarities.
- Types of clustering:
  - Connectivity based clustering

    objects related to nearby objects than to objects farther away
  - Centroid based clustering

    Cluster points according to a set of given centers
  - Distribution based clustering

    objects belonging most likely to the same distribution
  - Density based clustering

    areas of higher density than the remainder of the data set

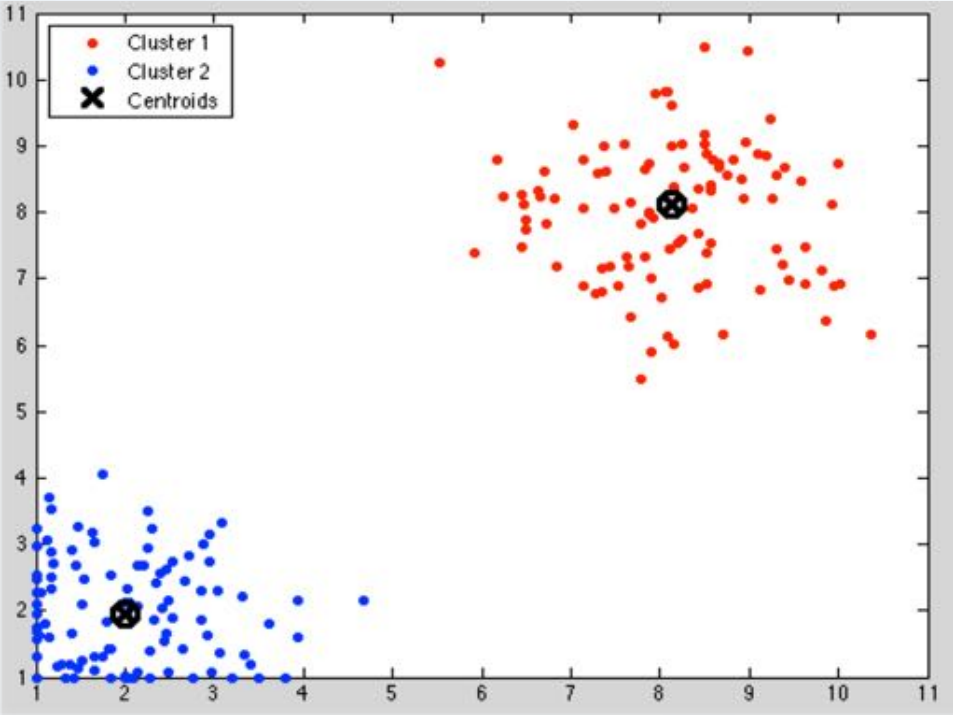


Unsupervised Learning

Clustering

Toy Example: A survey that has the following questions on a scale 1-10:

- How much do you like shopping?
- How much are you willing to spend on shopping?



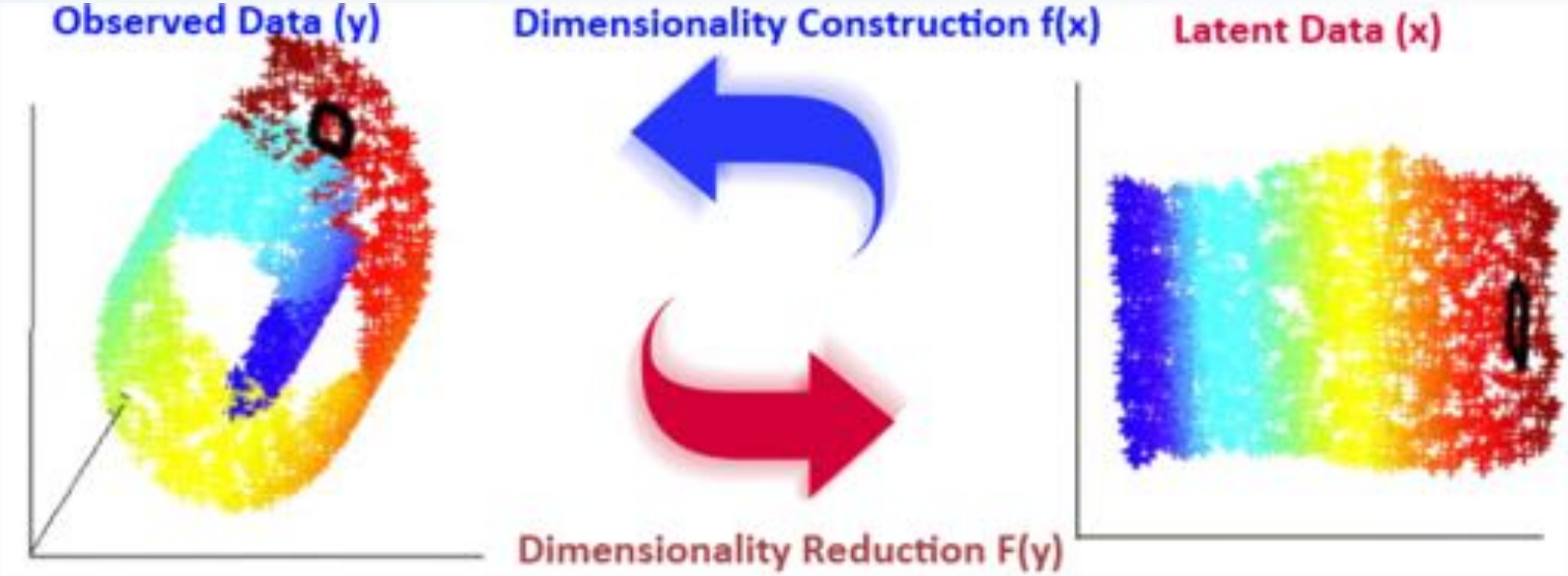
Cluster 1 can refer to people who are addicted to shopping

Cluster 2 can refer to people who rarely go shopping

Unsupervised Learning

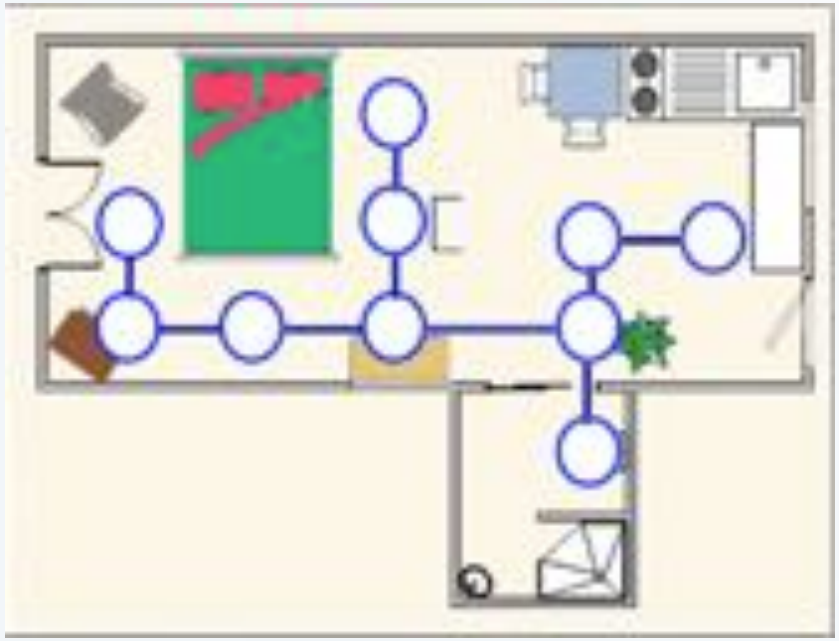
Dimensionality Reduction [7]

- Convert high dimensional data to lower order dimension
- Motivation:
  - High Dimensional Data Analysis
  - Visualization of high-dimensional data
  - Feature Extraction

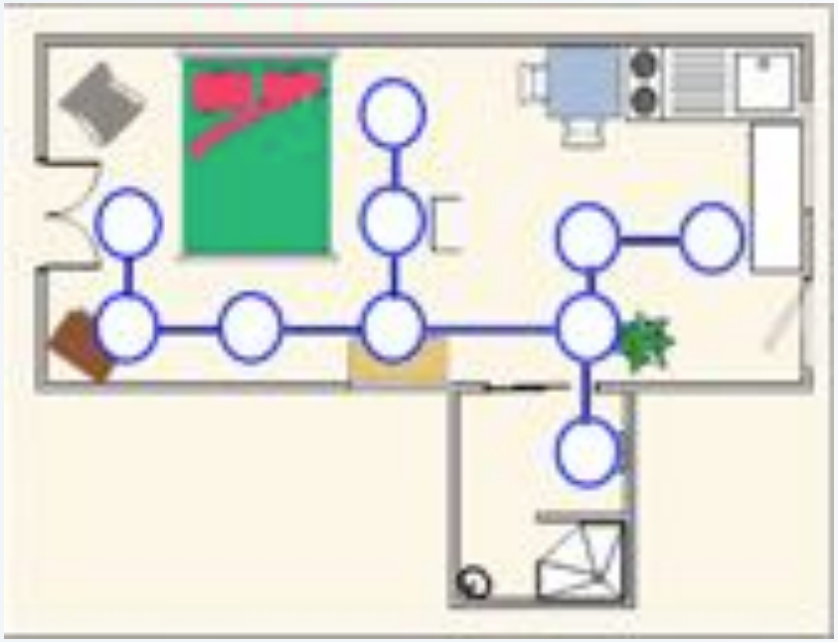


Reinforcement Learning

- Learning a policy: a sequence of outputs [1].
- Delayed reward instead of supervised output.
- Toy Example: A robot wants to move from the outer door of an apartment to the bathroom to clean it.

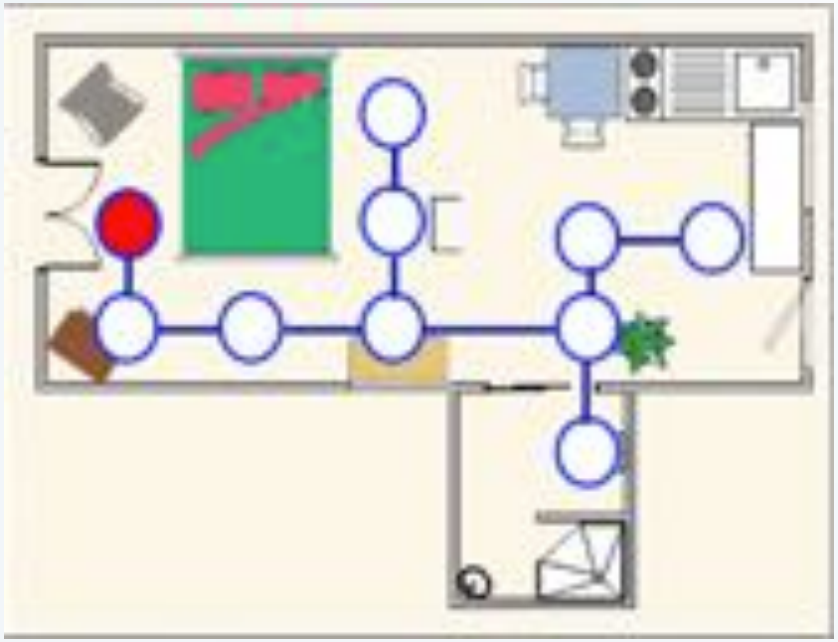


Reinforcement Learning

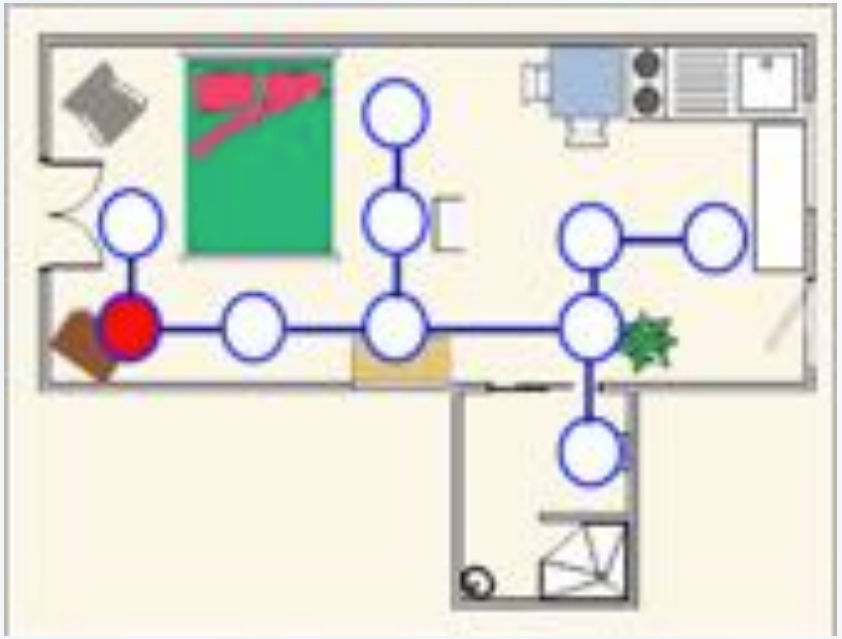


All weights are equal at the first try. Choice of next state is randomly chosen if the weights are equal

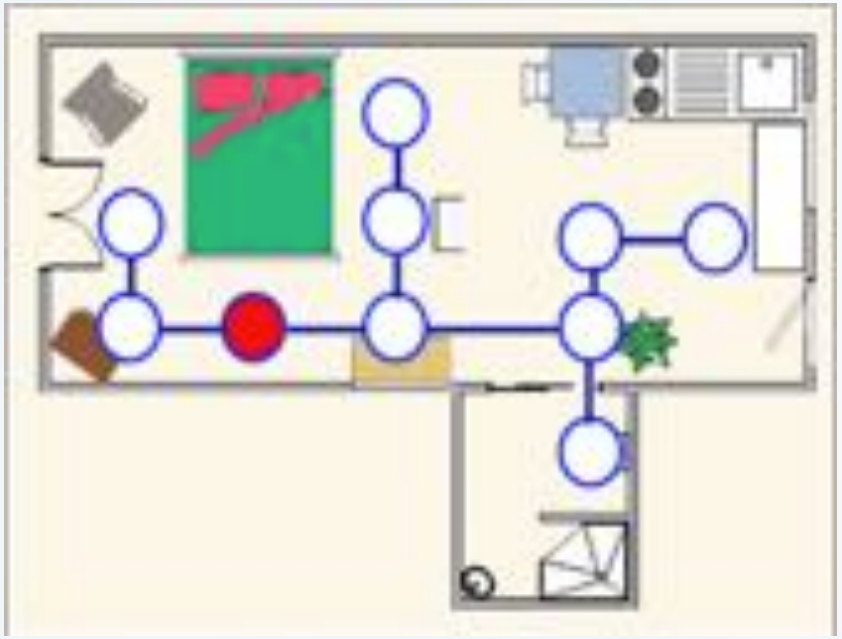
Reinforcement Learning



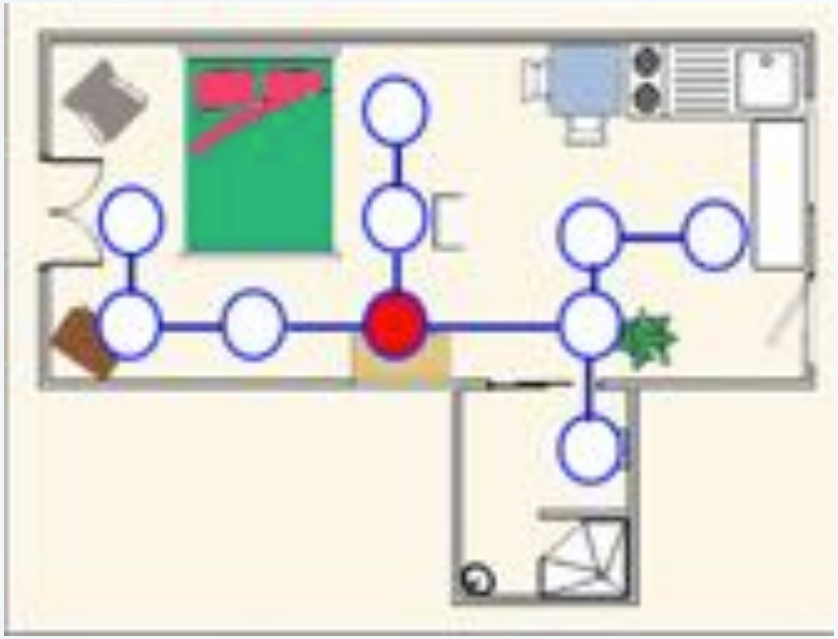
Reinforcement Learning



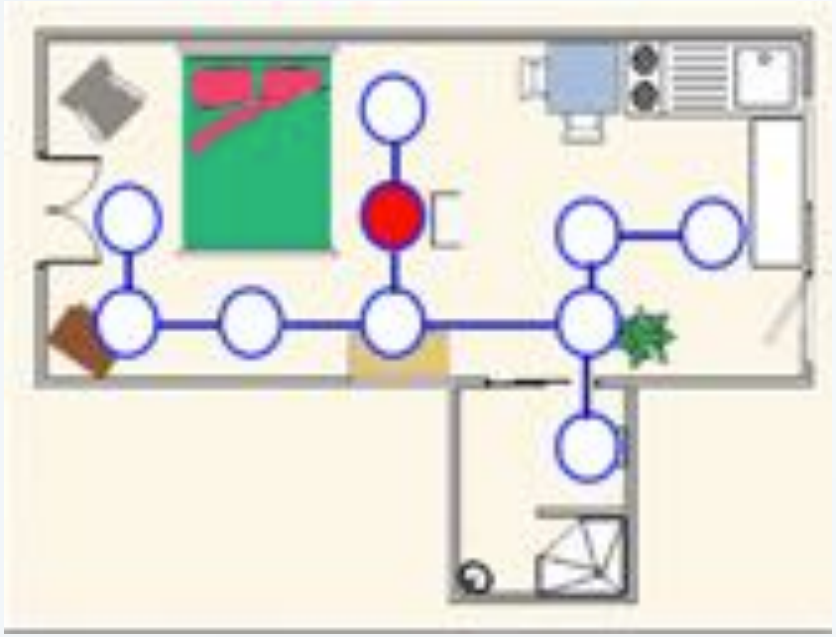
Reinforcement Learning



Reinforcement Learning

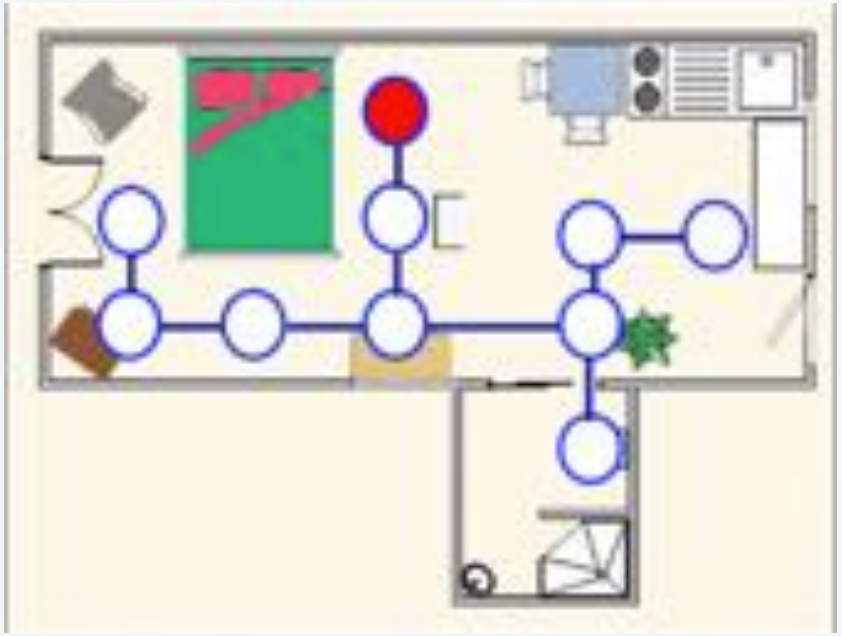


Reinforcement Learning



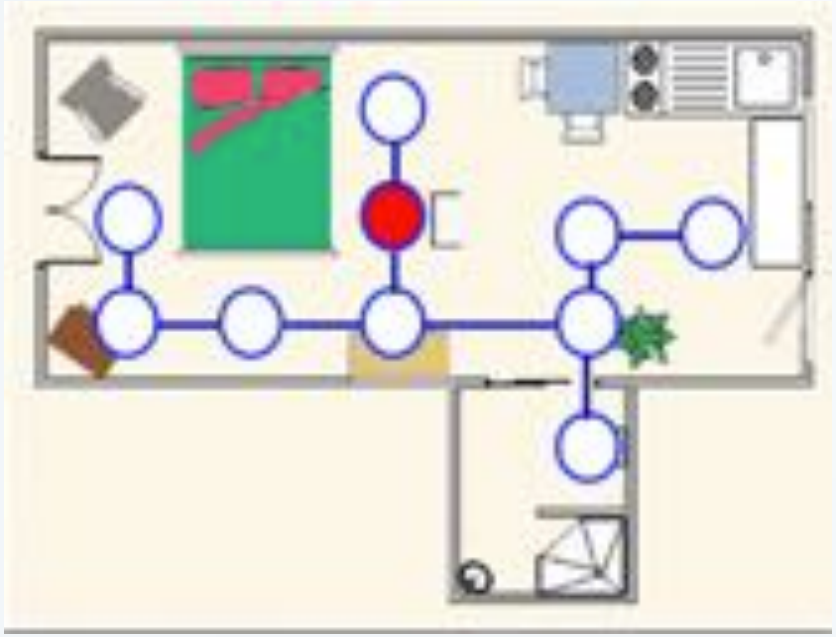
Left is chosen randomly since the weights are equal

Reinforcement Learning

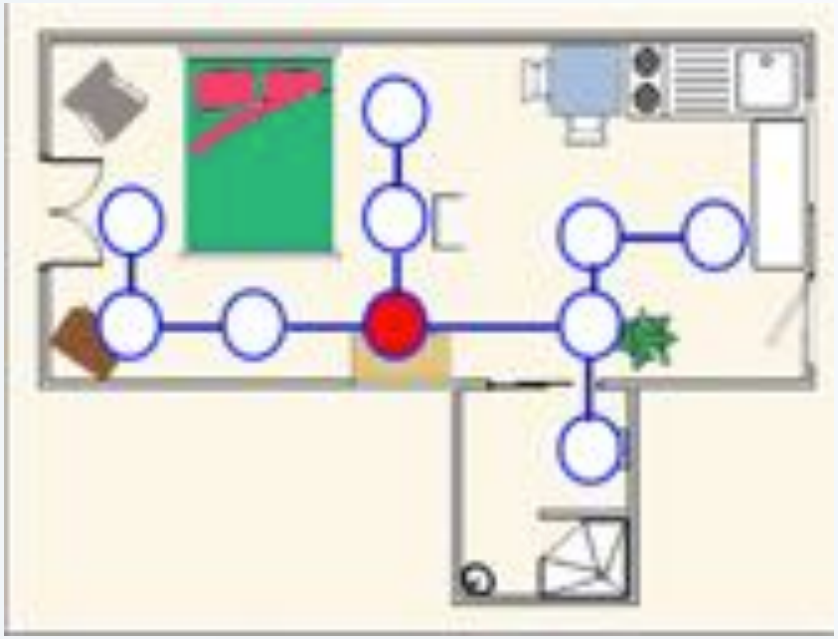


Wrong Destination.
Return by backtracking

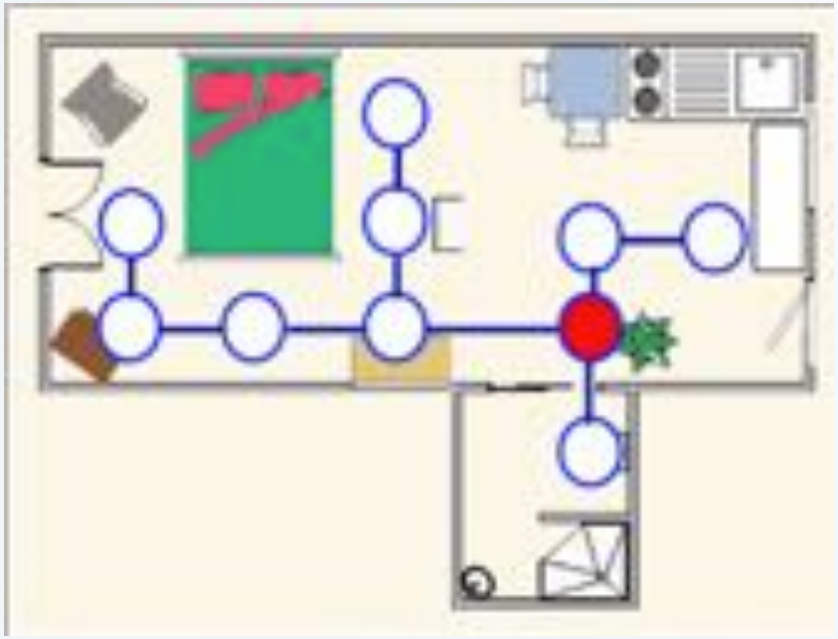
Reinforcement Learning



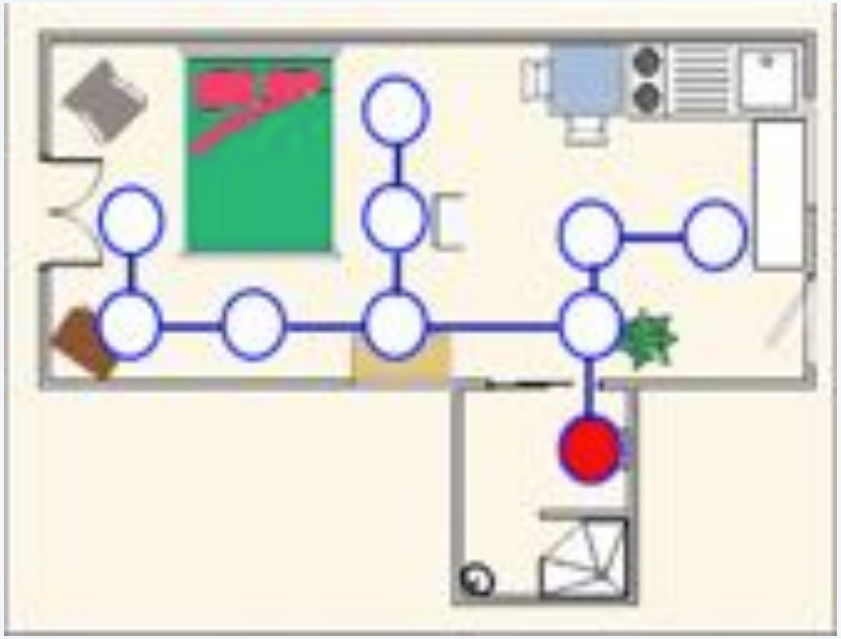
Reinforcement Learning



Reinforcement Learning

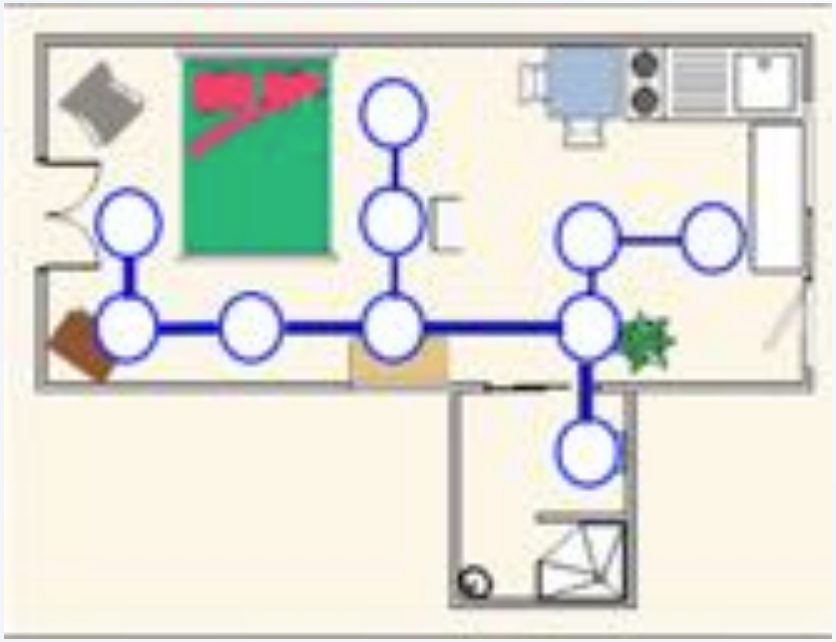


Reinforcement Learning



Reached the destination. Give a reward to the chosen paths by increasing the weight.

Reinforcement Learning



Adjusted weights after reinforcement learning.

Other Learning Paradigms

- Semi-Supervised Learning (Wikipedia)
- Active Learning (Wikipedia)
- Inductive Transfer/Learning (Wikipedia)

Real World Examples

Machine Learning in Real-World Examples: [6]

- Spam Filter
- Signature Recognition
- Credit Card Fraud Detection
- Face Recognition
- Text Recognition
- Speech Recognition
- Speaker Recognition
- Weather Prediction
- Stock Market Analysis
- Advertisement Targeting
- Language Translation
- Recommendation Systems
- Classifying DNA Sequences
- Automatic vehicle Navigation
- Object Detection
- Medical Diagnosis

Online Courses and Material

- Interactive Course with Stanford University Professor
  - Website: https://www.coursera.org/course/ml

- Stanford University Class
  - Playlist: http://www.youtube.com/view_play_list?p=A89DCFA6ADACE599
  - Material: http://cs229.stanford.edu/

References

1. E. Alpaydin, Introduction to Machine Learning. Cambridge, MA: MIT Press, 2004.
2. T. Mitchell, Machine Learning. McGraw Hill, 1997.
3. Liu H, Han J, Xin D, Shao Z (2006) Mining frequent patterns on very high dimensional data: a top-down row enumeration approach. In: Proceeding of the 2006 SIAM international conference on data mining (SDM'06), Bethesda, MD, pp 280–291.
4. C. Bishop. Pattern Recognition and Machine Learning. Springer, 2006.
5. T. Runkler. Data Analytics. Springer, 2012.
6. http://www.cs.utah.edu/~piyush/teaching/23-8-slides.pdf
7. Carreira-Perpinan, M., Lu, Z.: Dimensionality Reduction by Unsupervised Learning Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference, 1895 -1902 13 June 2010 .
8. S. Russell and P. Norvig: *Artificial Intelligence: A Modern Approach* Prentice Hall, *2003, Second Edition*
9. Sami Ghawi, Mustafa Jarrar: Lecture Notes on Introduction to Machine Learning, Birzeit University, 2018
10. Mustafa Jarrar: Lecture Notes on Decision Tree Machine Learning, Birzeit University, 2018
11. Mustafa Jarrar: Lecture Notes on Linear Regression Machine Learning, Birzeit University, 2018