CENG 463
Machine Learning

Lecture 01 - Introduction

About the Course

Lecturer: Asst. Prof. Nesli Erdoğmuş
E-mail: neslierdogmus@iyte.edu.tr
Exams: One midterm (~30%) and one final (~40%)
Assignments/projects: There will be a few programming assignments (to be completed in Python). Assignments will be initiated in lab sessions where Python will be introduced.
Course material: Slides, assignments and grades will be posted on cms.iyte.edu.tr.
Cheating: You're expected to solve the questions in the assignments and exams individually. Any copying of assignments and answers in the exam will be penalized according to the university policy.

References

- Christopher Bishop, Pattern Recognition and Machine Learning. Springer, 2006.
- Ethem Alpaydın, Introduction to Machine Learning (2nd Edition). MIT Press, 2010.
- Ethem Alpaydın, Yapay Öğrenme (in Turkish). Boğaziçi University Press, 2011.
- Richard Duda, Peter Hart and David Stork, Pattern Classification (2nd Edition). John Wiley & Sons, 2001.
- Stanford University's online course by Andrew Ng: Introduction to Machine Learning
  https://www.coursera.org/course/ml
  There is an 60-70% overlap between that course and ours.

Machine Learning is Everywhere

- Spam e-mail filters learn to detect spam mails.
- Amazon uses a learning algorithm to recommend you new books based on the book you have just selected.
- Facebook uses face recognition algorithms for auto-tagging.
- US postal service automatically recognizes handwriting to read postal codes.
- **Closely related to:**
  - Data mining
  - AI
  - Pattern Recognition
  - Robotics
  - Computer Vision

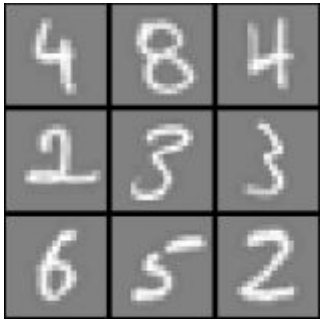
Good to Know

This course requires some background in:

- Linear Algebra
  - Matrices, vectors and operations of these
  - Eigenvalues and Eigenvectors
  - Singular Value Decomposition
- Statistics
  - Basic concepts (mean, variance etc.)
  - Gaussian Distribution
  - Bayesian Rule
- Optimization
  - Gradient descent, learning rate etc.

What is Machine Learning?

- Arthur Samuel (1959): Machine Learning can be defined as ‘the field of study that gives computers the ability to learn without being explicitly programmed’.
- Calculating tax returns for example, is not learning!
- It is learning if we only say what to / how to learn.
  - Example: Computing the location of white pixels of handwritten numbers and learning the distribution within the square.



What is Machine Learning?

- Tom Mitchell (1998) Well-posed Learning Problem:

"A computer program is said to **learn** from **experience E** with respect to some **task T** and some **performance measure P**, if its performance on **T**, as measured by **P**, improves with experience **E**."

What is Machine Learning?

“A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.”

Example: Spam filter

- T is classifying emails as spam and non-spam.
- P is the ratio of detected spams / real spams.
- E is user’s feedback.
- If with added experience E, its performance P on task T increases, then it is learning!

What is T, P and E in Amazon book recommender?

Why we make computers learn?

Learning is used when:

- Human expertise does not exist or very expensive due to the huge amount of data (Amazon recommender)
- Humans are unable to explain their expertise (speech recognition, object recognition)
- Solution changes in time (routing on a computer network)
- Solution needs to be adapted to particular cases (user biometrics)

Types of Learning

Machine learning algorithms:

- **Supervised learning:**
  The ‘right’ answers are given while learning.

- **Unsupervised learning:**
  There is no ‘right’ or ‘wrong’ answer, there are groups/clusters to be identified.

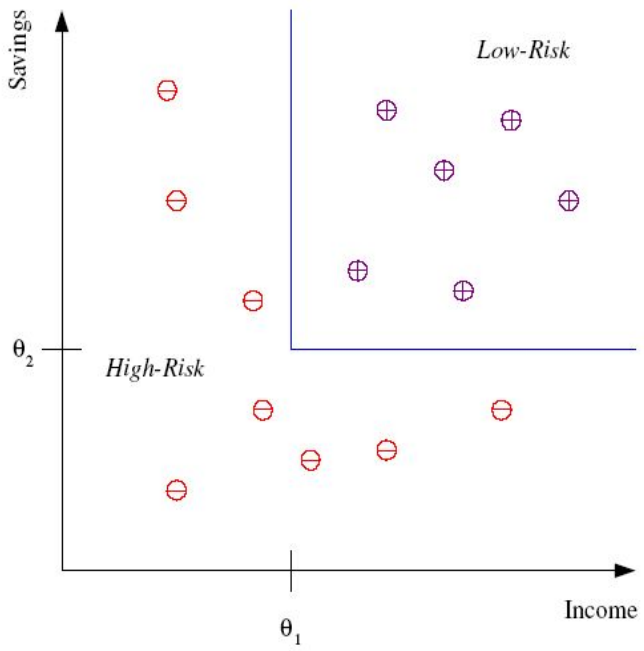
- **Reinforcement learning:**
  There is no ‘right’ or ‘wrong’ answer, but the machine learns to make right decisions with a reward-penalty system.

Supervised Learning: Classification

Example: Bank credit scoring

- Differentiating between low-risk and high-risk customers based on their *income* and *savings*.
- There are two classes to be distinguished from each other.This is called **classification**.
- Classifier has two parameters: θ_1 and θ_2

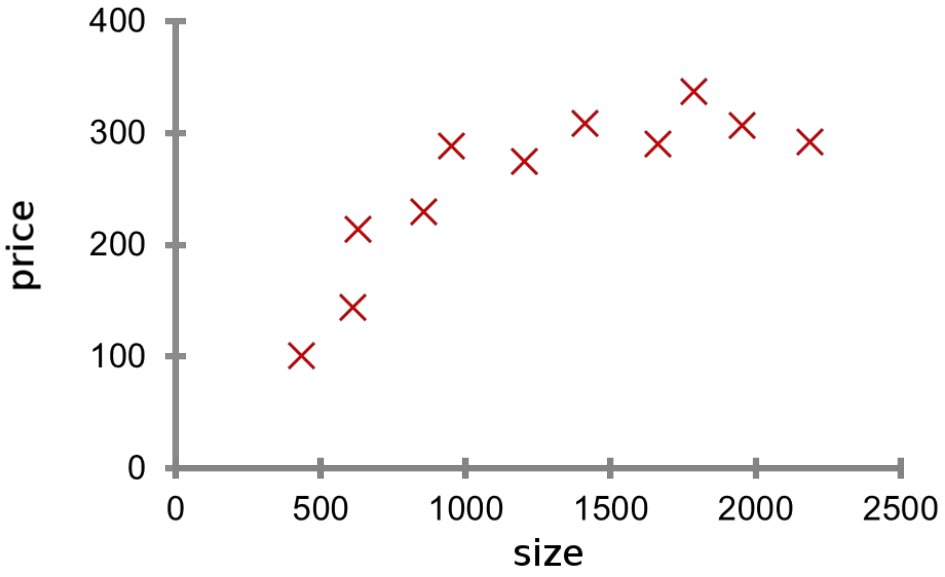
IF income > θ_1 **AND** savings > θ_2:
THEN: low-risk
ELSE: high-risk



Supervised Learning: Regression

Example: Housing price prediction

- There are no classes but there is a continuous valued output (*price*) with respect to an attribute (*size*). This is called **regression**.



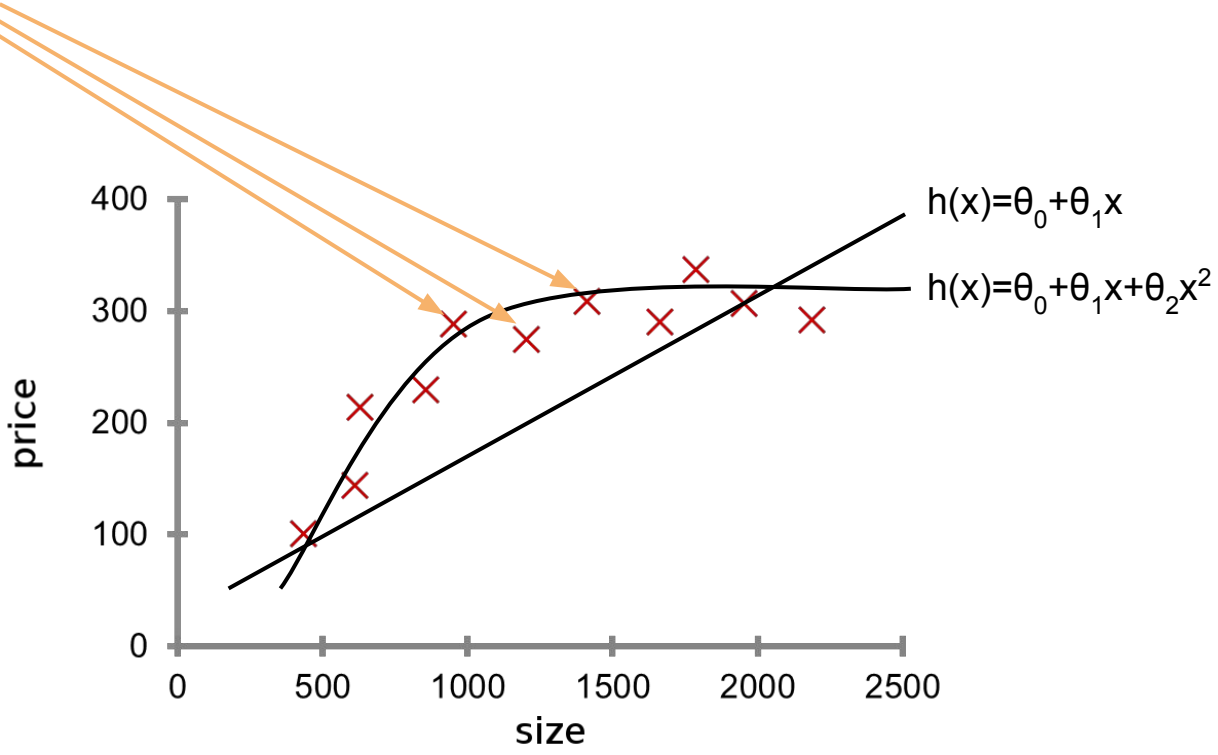
Classification or Regression?

Should you treat these as classification or as regression problems?

- **Problem 1:** You have a large stock of a certain product in your company. You want to predict how many items will be sold next month.
- **Problem 2:** Spam e-mail detection.
- **Problem 3:** Navigating a car, estimating the angle of the steering.
- **Problem 4:** Deciding if a person is male or female given his/her photo?

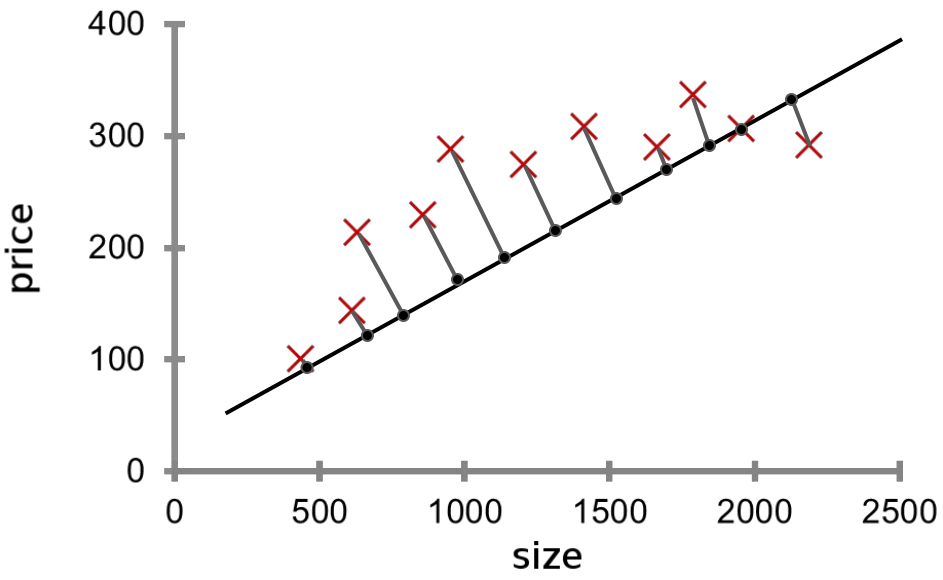
Model Selection

We need a model, for which the parameters are learned from the training samples.



Cost Function

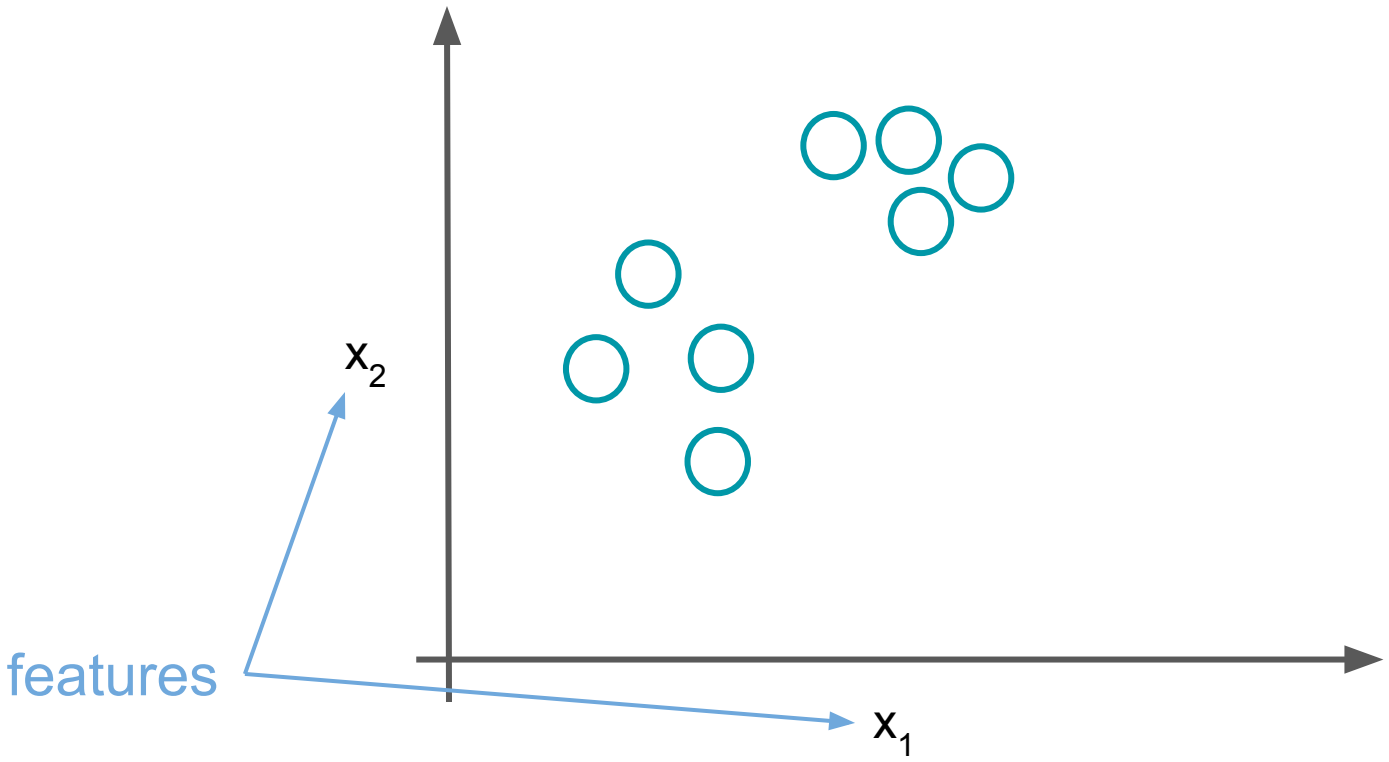
To find the parameters of our model, we need a cost function to minimize.



- y: real prices (red crosses)
- Our hypothesis: $h(x)=\theta_0+\theta_1 x$
- First sample cost: $(y^{(1)} - h(x^{(1)}))^2$
- Total cost: $\Sigma(y^{(i)} - h(x^{(i)}))^2$

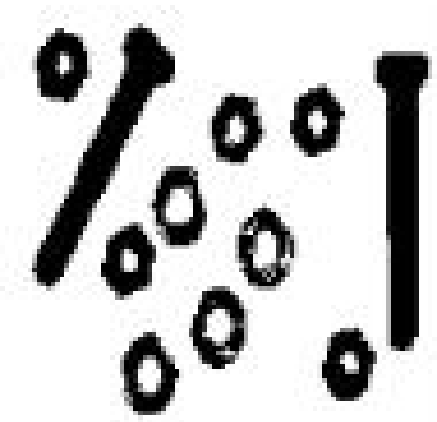
Unsupervised Learning

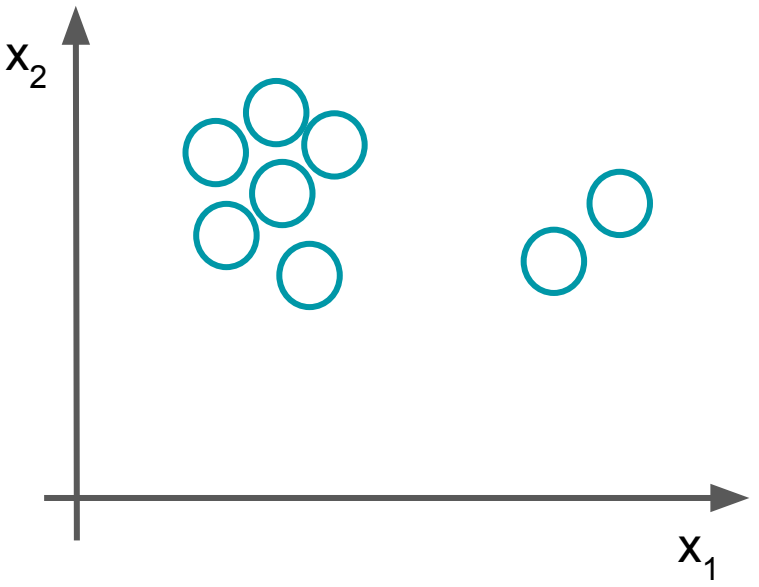
Now, we have different types of instances but we do not have 'labels'. We cluster (group) the samples using some features.



Unsupervised Learning

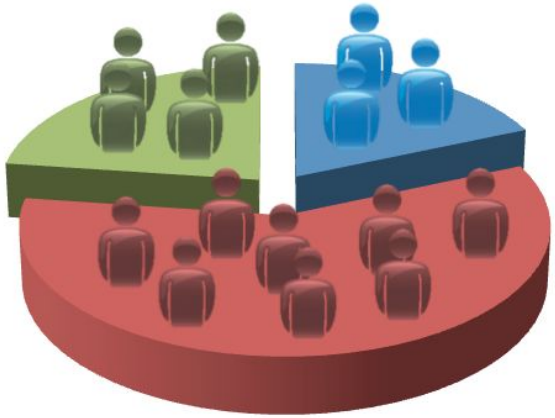
Example: How many different types of parts?



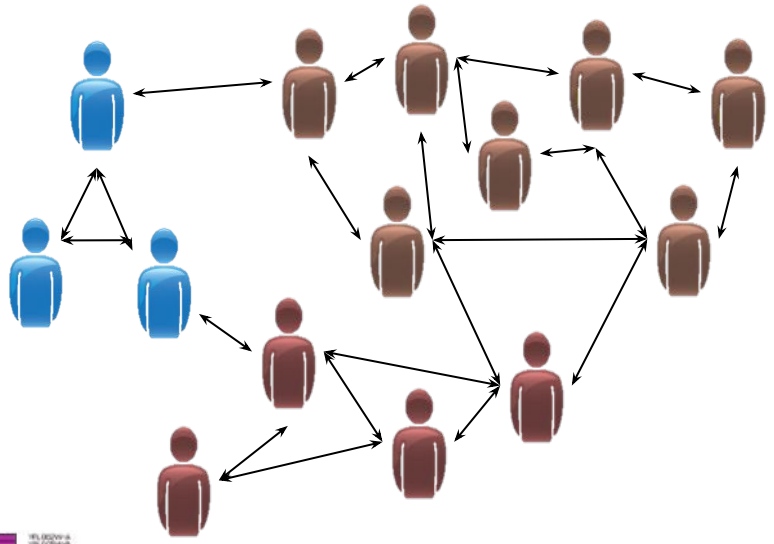


What may be the features x_1 and x_2?

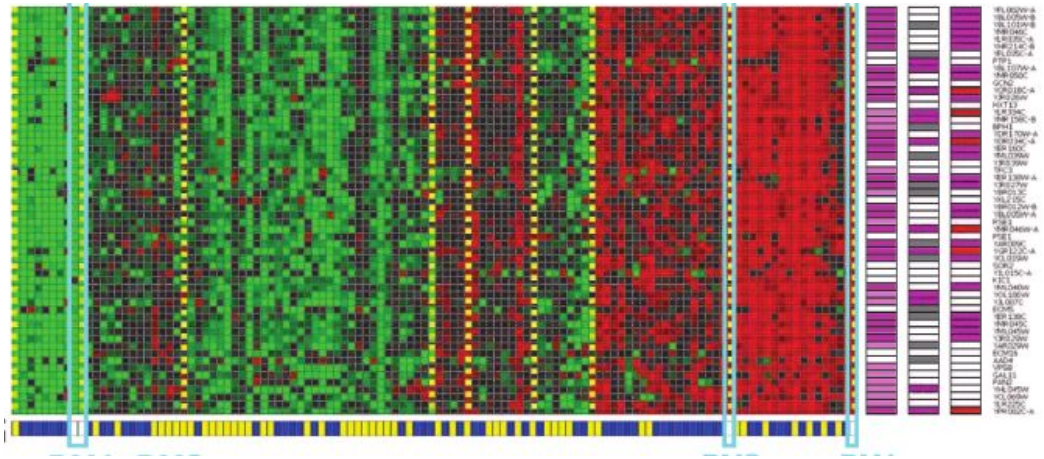
Unsupervised Learning



Market segment analysis



Social network analysis



Bioinformatics: Learning motifs

Unsupervised Learning

Supervised or Unsupervised?

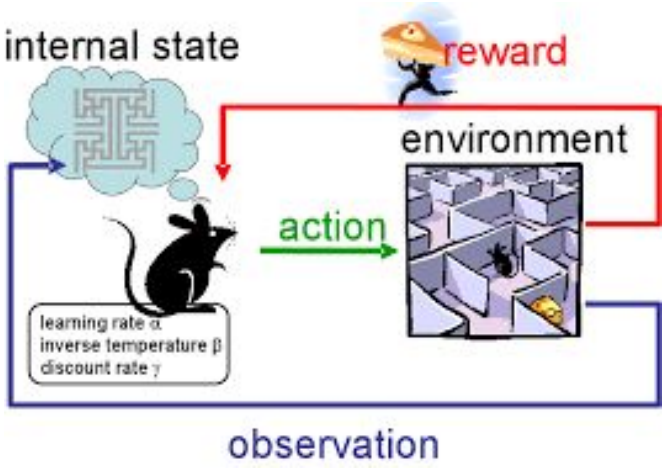
Should you use a supervised or unsupervised learning algorithm?

- **Problem 1:** Given email labeled as spam/not spam, learn a spam filter.
- **Problem 2:** Given a set of news articles found on the web, group them into sets of articles about the same story.
- **Problem 3:** Given a database of customer data, automatically discover market segments and group customers into these segments.
- **Problem 4:** Given a dataset of patients diagnosed as either having diabetes or not, learn to classify new patients as having diabetes or not.

Reinforcement Learning

Reinforcement learning is learning what to do - how to map situations to actions- so as to maximize a numerical reward signal.

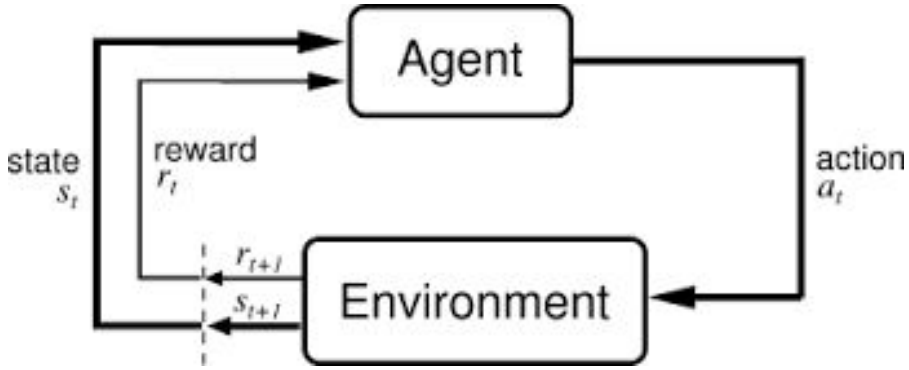
- **Learner:** Agent
- **Environment:** States
- **What to do:** Action
- **Result:** Reward or punishment



The environment is a modeled as a stochastic finite state machine with inputs (actions sent from the agent) and outputs (observations and rewards sent to the agent)

Reinforcement Learning

- The basic idea is simply to capture the most important aspects of the real problem facing a learning agent interacting with its environment to achieve a goal.
- Reinforcement learning is different from supervised learning, that is learning from examples provided by a knowledgeable external supervisor. In reinforcement learning, the agent learns through interactions.



Reinforcement Learning

Trade-off between exploration and exploitation:

● To obtain a lot of reward, a reinforcement learning agent must prefer actions that it has tried in the past and found to be effective in producing reward.

● But to discover such actions, it has to try actions that it has not selected before.

→ The agent has to **exploit** what it already knows in order to obtain reward, but it also has to **explore** in order to make better action selections in the future.