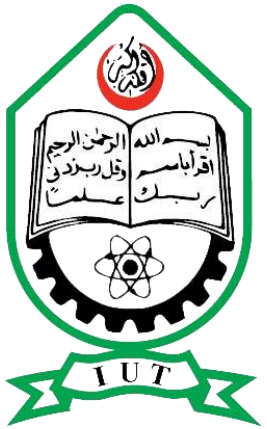


CSE 4621
Machine Learning

Lecture 1

Md. Hasanul Kabir, PhD.

Professor, CSE Department

Email: hasanul@iut-dhaka.edu

Introduction

- "Data is abundant and cheap but knowledge is scarce and expensive."

- Machine: Computer / Computer Program
- Learning: *ability to learn without being explicitly programmed.*

- 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.*

Examples

❑ **A checkers learning problem:**

- **Task T:** playing checkers
- **Performance measure P:** percent of games won against opponents
- **Training experience E:** playing practice games against itself

❑ **A handwriting recognition learning problem:**

- **Task T:** recognizing and classifying handwritten words within images
- **Performance measure *P:*** percent of words correctly classified
- **Training experience E:** a database of handwritten words with given classifications

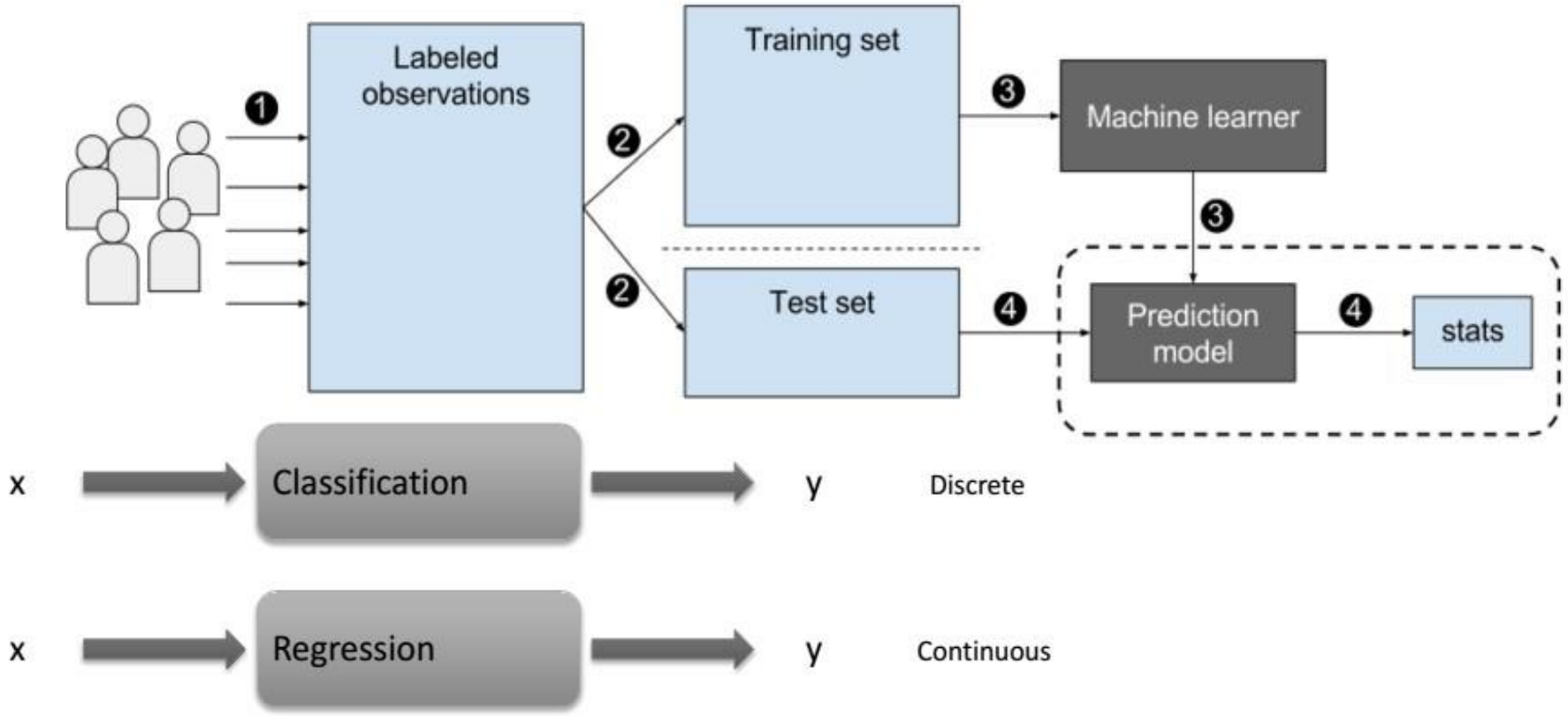
❑ **A robot driving learning problem:**

- **Task T:** driving on public four-lane highways using vision sensors
- **Performance measure P:** average distance traveled before an error (as judged by human overseer)
- **Training experience E:** a sequence of images and steering commands recorded while observing a human driver

Types of Learning

- **Supervised** (inductive) learning
  - Training data includes desired outputs
- **Unsupervised** learning
  - Training data does not include desired outputs
- **Semi-supervised** learning
  - Training data includes a few desired outputs
- **Reinforcement** learning
  - Rewards from sequence of actions
- **Active** learning
  - Let users play an active role in the learning process.

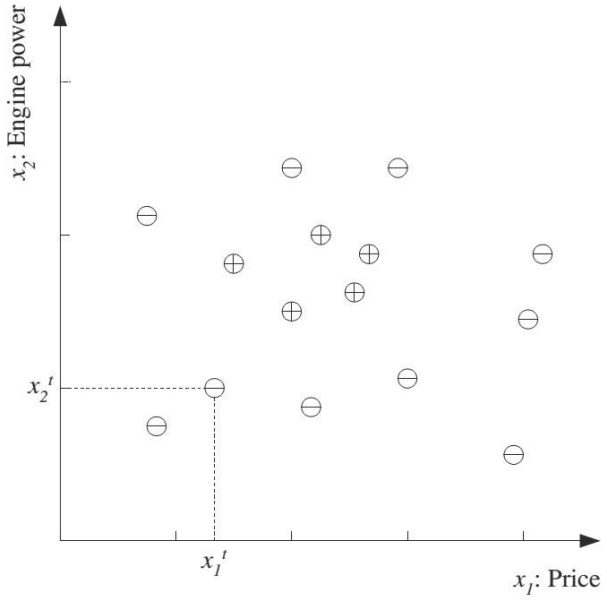
Supervised Learning

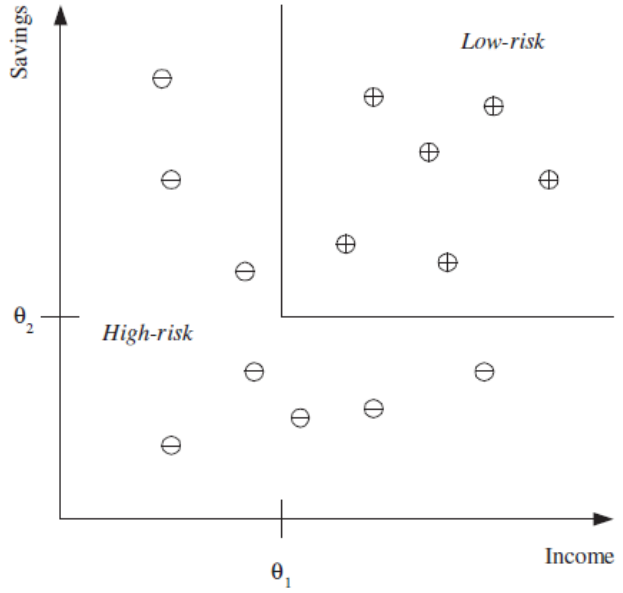


- Common Learning Algorithms: Linear Regression, Logistic Regression, Decision Tree, Random Forest, k-Nearest Neighbour, SVM, Neural Network etc.

Supervised Learning

- Given a set of data points $\{x^1, x^{(2)}, .., x^{(m)}\}$ associated to a set of outcomes, $\{y^1, y^{(2)}, .., y^{(m)}\}$, we want to build a classifier that learns how to predict y from x.





- Estimate the prediction model/function h, by minimizing loss function, e.g., $(y^{(i)} - h(x^i))^2$

Unsupervised Learning

- Looks for previously undetected patterns in a data set with no pre-existing labels.
  - Actual output y is absent!

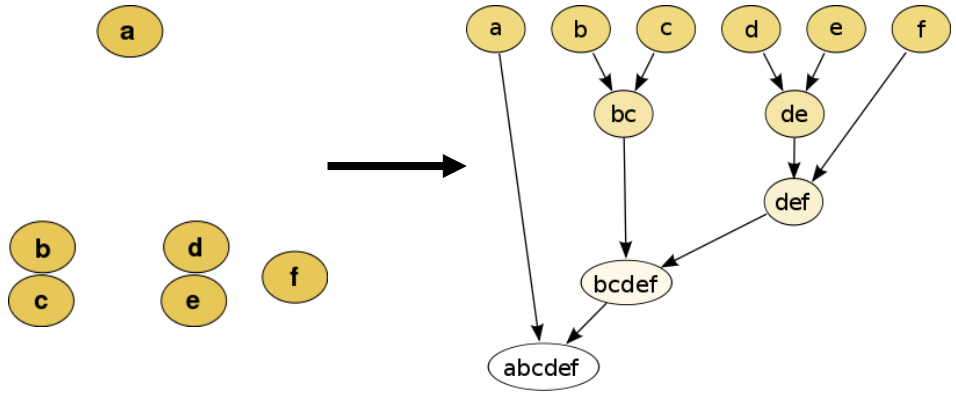


Fig: Agglomerative clustering

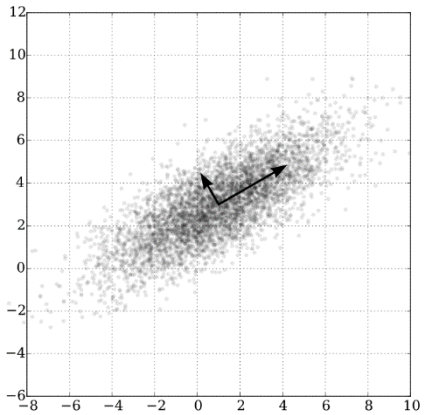


Fig: PCA

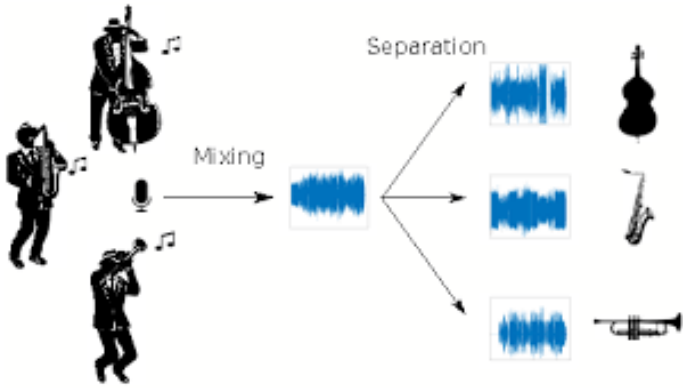
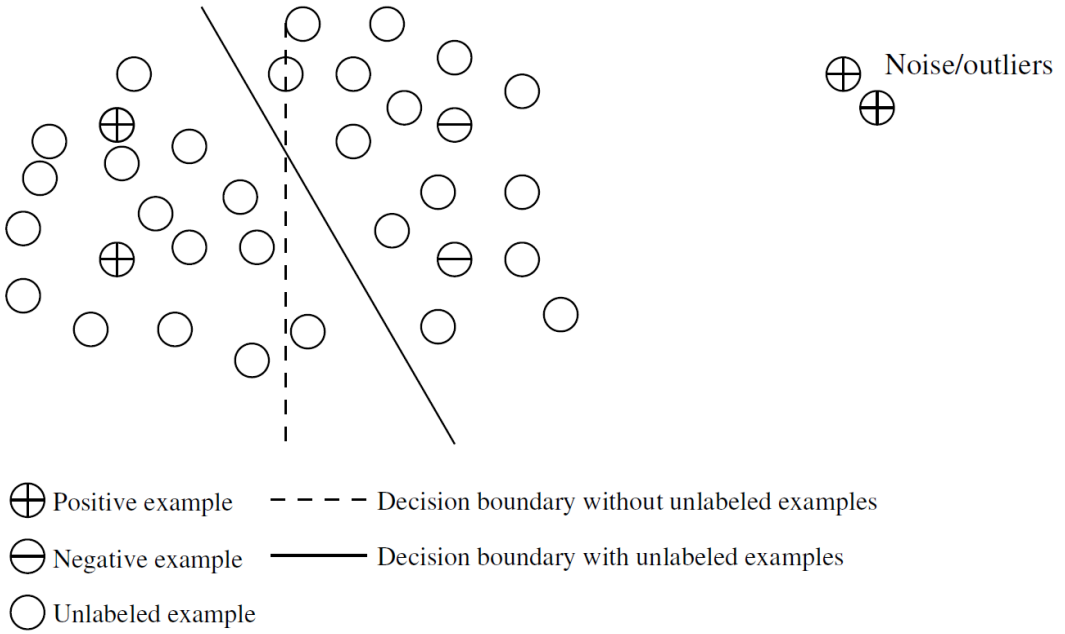


Fig: ICA

- Common Learning Algorithms: K-means, hierarchical clustering, mixture model, Local outlier factor, EM, PCA, ICA, etc.

Semi-Supervised Learning

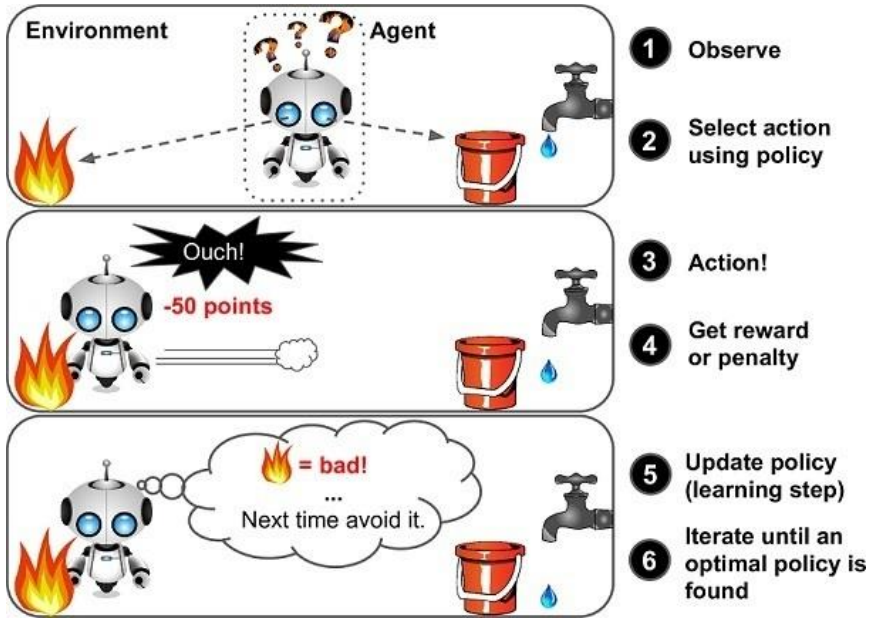
- Makes use of both labeled and unlabeled examples when learning a model.
- Labeled examples are used to learn class models and unlabeled examples are used to refine the boundaries between classes.

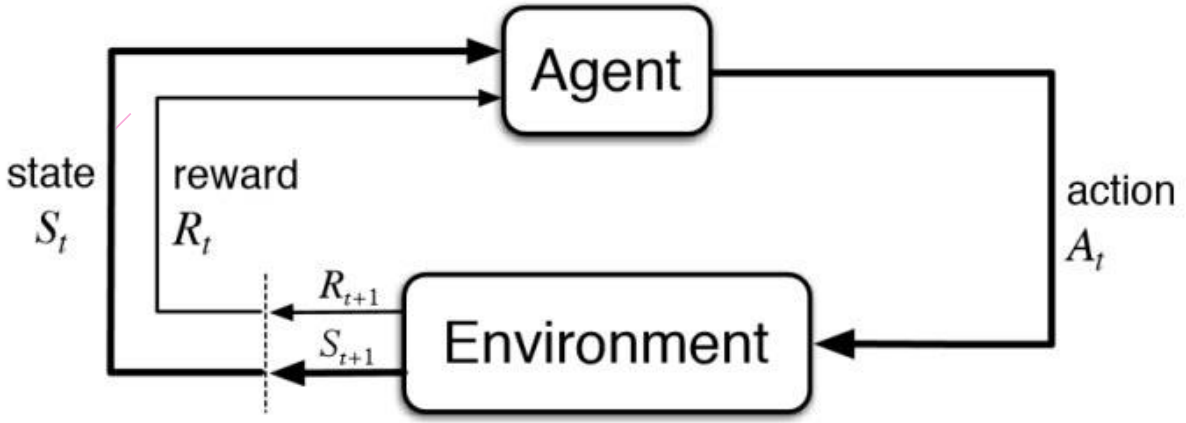


Semi-supervised learning.

Reinforcement Learning

- Should be able to assess the goodness of policies and learn from past good action sequences to be able to generate a policy.
  - what is important is the *policy* that is the sequence of correct actions to reach the goal.





- Common Learning Algorithms: Markov Decision Process, Q-Learning, Deep Q Learning, etc.

Basic Steps to Machine Learning

Our Focus

- Batch learning vs. Online learning
  - Batch: All data available before training
  - Online: Data samples come one after another
- Passive learning vs. Active learning
  - Passive learning: Only observing given data
  - Active learning: Can ask labels for some data as the learning machine wants
- Structured prediction
  - y itself has internal structures, e.g., a sequence, a tree.
- This course mainly focuses on supervised, batch, passive, non-structured ML