7. Machine Learning

7.1. Introduction to Learning by Rational Agents

✓ Learning: Improving performance through careful study of experiences.

✓Rational agents must learn, because
       'Designers neither can anticipate all possible future situations and changes over time, nor have idea of solutions for some problems'.

✓Various approaches to learning:

- Inductive versus Deductive Learning
- Empirical, statistical and example based learning
- Explanation based learning

✓Types of learning based on feedback types:

- Unsupervised, for example, partitioning of objects without given labeled examples

- Reinforcement, for example, fixing rules through rewards and punishments

- Supervised, for example, discovering rules using labeled training dataset (**commonly used**)

- Semi-supervised, Hypotheses verification, Generalization, etc.

✓ Learning models in practice:

Linear regression, nonlinear regression, decision tree, neural network, naïve Bayes, clustering, support vector machines, etc.

✓ Learning by ensembles of models

➢ First we take a simple example of supervised learning (example based, inductive).

➢ Most examples are taken from 'Data Mining: Concepts and Techniques, 3rd Edition, Morgan Kaufmann Publishers, 2012, by Jiawei Han, Micheline Kamber & Jian Pei'.

➢ In general, the task of supervised learning is this:

Given a **training set** of N example input-output pairs
$(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)$,
where each y_i was generated by an unknown function $y=f(x)$,
discover a function h that approximates the function f.

x can be numeric or non numeric; h is a hypothesis from a hypothesis space H.

7.2 Learning Decision Trees

➢Training Samples: [Described through attribute values along with the class they belong to]

ID	Age	Income	Student	Credit Rating	Decision/ Class/ Label
1	≤ 30	high	no	fair	negative
2	≤ 30	high	no	excellent	negative
3	31…40	high	no	fair	positive
4	> 40	medium	no	fair	positive
5	> 40	low	yes	fair	positive
6	> 40	low	yes	excellent	negative
7	31…40	low	yes	excellent	positive
8	≤ 30	Medium	no	fair	negative
9	≤ 30	low	yes	fair	positive
10	> 40	medium	yes	fair	positive
11	≤ 30	medium	yes	excellent	positive
12	31…40	medium	no	excellent	positive
13	31…40	high	yes	fair	positive
14	> 40	medium	no	excellent	negative

Class Concepts: positive ('Buys a computer', and negative ('Does not buy a computer')

Some Rules, such as 'A student of age not over 30 buys a computer', ultimately show up that can classify an unknown sample.