CS434
Machine Learning and Data Mining

Fall 2008

Administrative Trivia

- Instructor:
  - Dr. Xiaoli Fern (Back on Wednesday)
  - web.engr.oregonstate.edu/~xfern
  - Office hour: 1 hour before class, or by appointment
- Course webpage

  web.engr.oregonstate.edu/~xfern/classes/cs434
- Please check course webpage frequently
  - Learning objectives
  - Syllabus
  - Course policy
  - Course announcements

Briefly

- Grading:
  - Homeworks and projects – 55%
  - Midterm – 20%
  - Final exam – 25%
- Homeworks
  - due at the beginning of the class (first 5 minutes of the class)
  - Late submission will be accepted if it's no more than 24 hours late, but only gets 80%
- Collaborations policy (for solo assignments)
  - Verbal discussion about general approaches and strategies allowed
  - Can talk about examples not in the assignments
  - Anything you turn in has be created by you and you alone

For team assignments, the above policies apply between teams.

Course materials

- No text book required, slides and reading materials will be provided on course webpage
- There are a number of recommended books that are good references
  - Machine learning by Tom Mitchell (TM)
  - Pattern recognition and machine learning by Chris Bishop (Bishop)

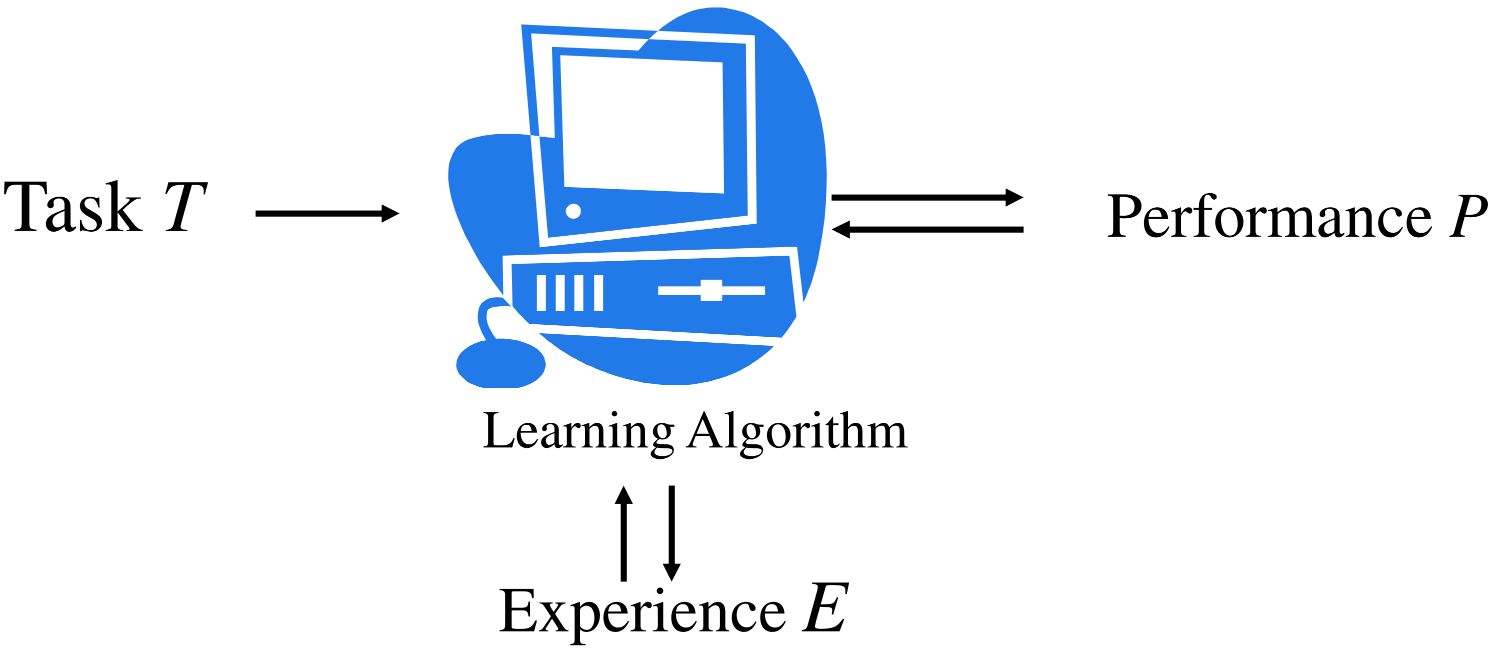
What is learning?

Generally speaking

“any change in a system that allows it to perform better the second time on repetition of the same task or on another task drawn from the same distribution”

--- Herbert Simon

Machine learning



Task T

Learning Algorithm

Performance P

Experience E

Learning = Improving with experience at some task

- Improve over task T
- with respect to P
- based on experience E

When do we need computer to learn?

What is not learning?

- × A program that does tax return
- × A program that looks up phone numbers in phone directory
- × …

When do we need learning?

- Sometimes there is no human expert knowledge
  - Predict whether a new compound will be effective for treating some disease
- Sometimes humans can do it but can’t describe how they do it
  - Recognize hand written digits
- Sometimes the things we need to learn change frequently
  - Stock market, weather forecasting, computer network routing
- Sometimes the thing we need to learn needs customization
  - Spam filters

Fields of Interest

- Supervised learning – learn to predict
- Unsupervised learning – learn to understand and describe the data
- Reinforcement learning – learn to act

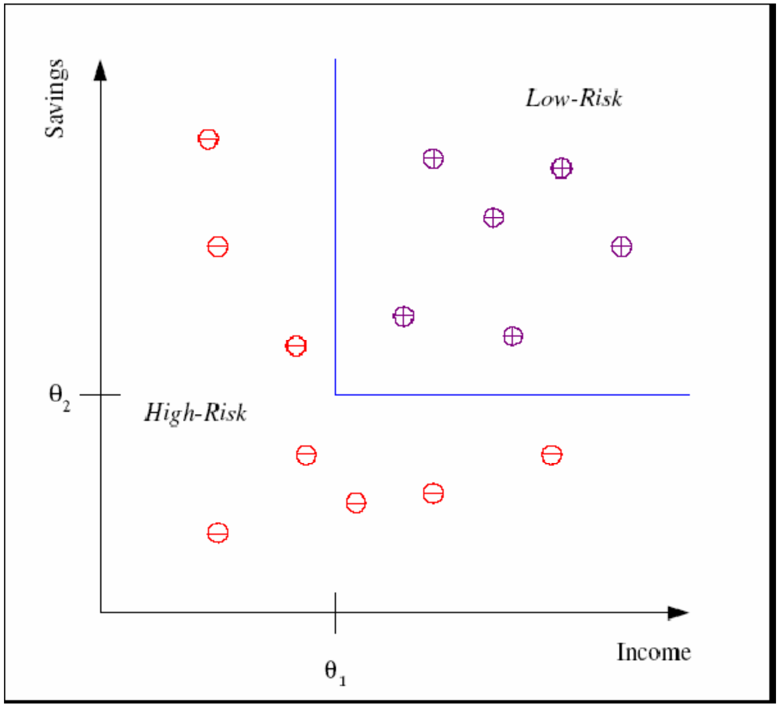
Data mining

A highly overlapping concept, but focuses on large volume of data:

To obtain useful knowledge from large volume of data

Supervised Learning: example

- Learn to predict output from input
  - E.g. predict the risk level of a loan applicant based on income and savings



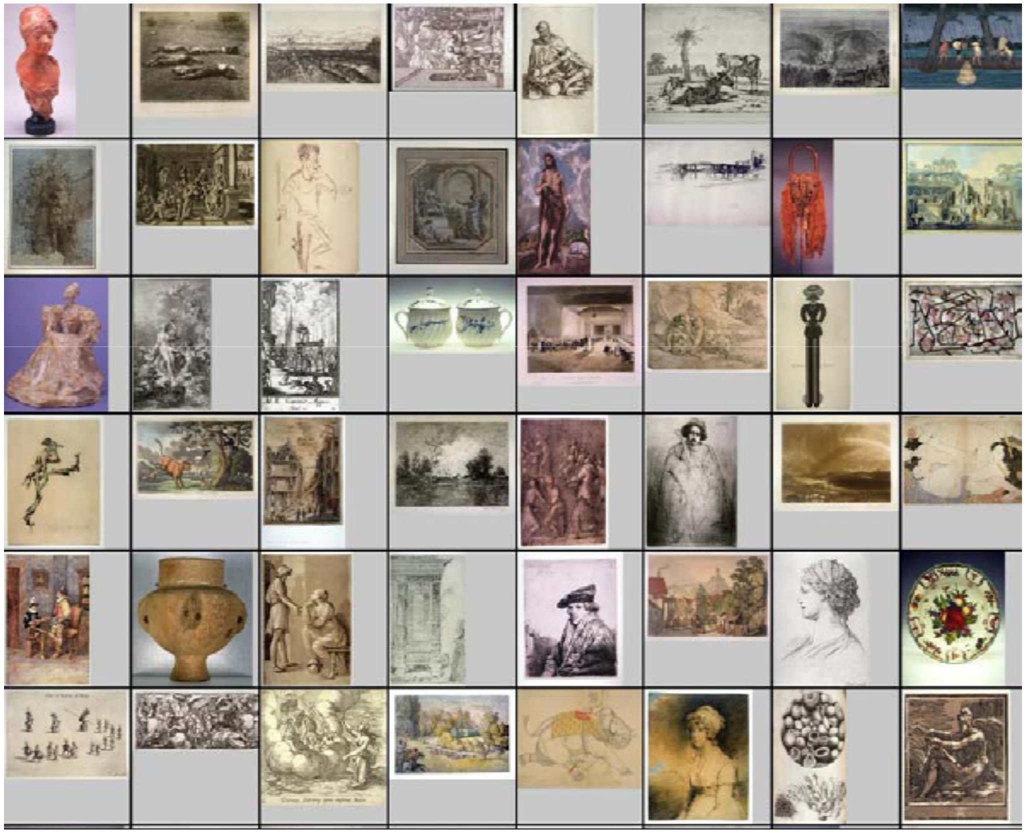
MANY interesting applications!

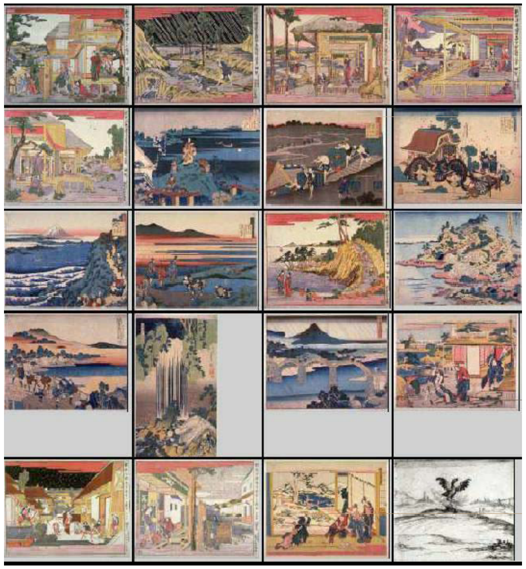
Spam filters,

Collaborative filtering (predicting if a customer will be interested in an advertisement),
Ecological (predicting if a species is absent/present in a certain environment),
Medical ……

Unsupervised learning

- Find patterns and structure in data



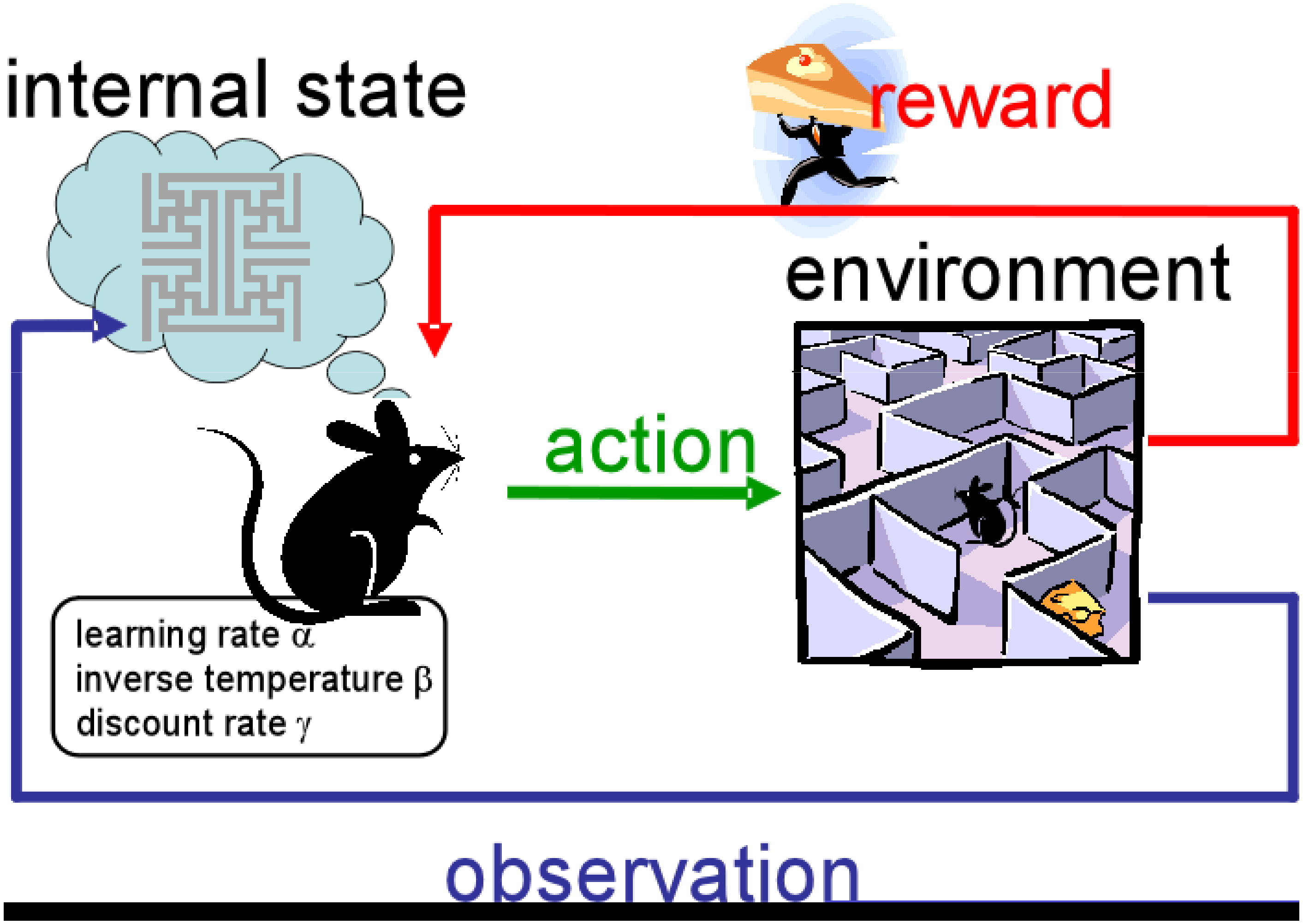


Clustering art

Example Applications

- Market Segmentation: divide a market into distinct subsets of customers
  - Collect different attributes of customers based on their geographical and lifestyle
  - Find clusters of similar customers, where each cluster may conceivably be selected as a market target to be reached with a distinct marketing strategy
- Document clustering
  - For organizing search results etc.

Reinforcement learning



Example Applications

- Robot controls
- Elevator scheduling
- Games such as backgammon and chess
- …

Learning objectives

- Students are able to ***apply supervised learning algorithms to prediction problems and evaluate the results.***
- Students are able to ***apply unsupervised learning algorithms to data analysis problems and evaluate results.***
- Students are able to ***apply reinforcement learning algorithms to control problem and evaluate results.***
- Students are able to ***take a description*** of a new problem and ***decide what kind of problem*** (supervised, unsupervised, or reinforcement) it is.

Example: Learning to play checkers

- *T:* play checkers
- *P*: percent of games won in world tournament
  - What experience?
  - What should we exactly learn?
  - How should we represent it?
  - What specific algorithm to learn it?

Type of training experience

- Direct
  - For each board state, we obtain a best move for that position
  - Observe many states and many moves
  - Try to learn what is the best move for an unseen state
- Indirect
  - Just observe a sequence of plays and the end result
  - More difficult, because
    - which of the moves are the bad (good) ones for a bad (good) game?
    - This is the credit assignment problem, very difficult to solve

Choose the Target Function (what should we learn)

- Choosemove: board state -> move?
- V: Board state -> Reward (value of the state)?
- …

Possible definition for target function V

- If b is a final board state that won, V(b)=100
- If b is a final board state that is lost, V(b)= -100
- If b is a final board state that is drawn, the V(b)=0
- If b is not a final board state, then V(b)=V(b’), where b’ is the best possible final state reachable from b.

This gives correct values, but is not operational

Choose representation for target function

- Collection of rules
- Neural network?
- Polynomial functions of board features?
- …

A representation for learned function

$$w_0 + w_1 f_1(b) + w_2 f_2(b) + \cdots + w_n f_n(b)$$

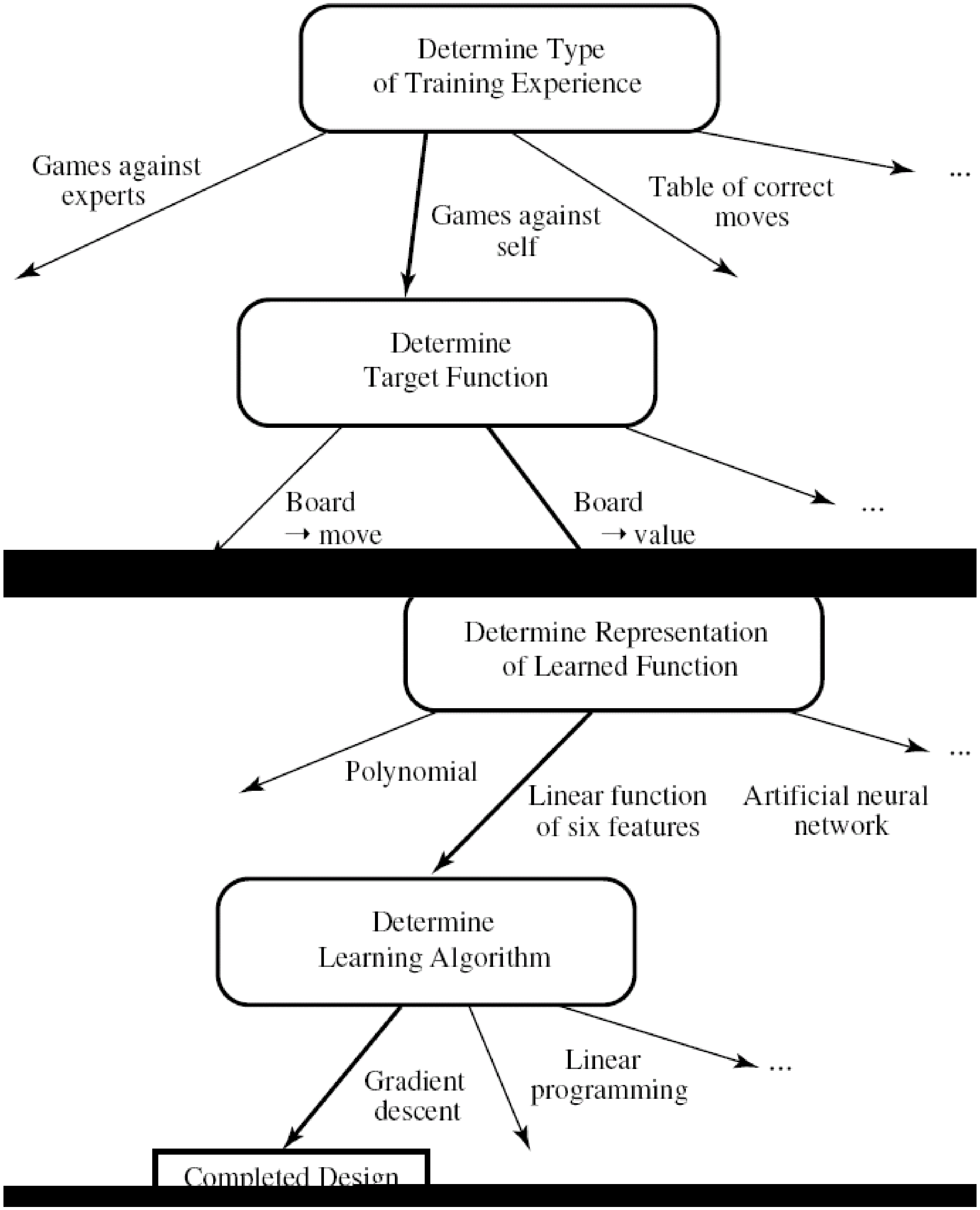
f1, f2, …, fn are features describing a board state

For example, f1 can be the number of black pieces on board
f2 can be the number of red pieces on board, etc.

A diagram of design choices

In this class, you will become familiar with many of these choices, and even try them in practice.

We would like to prepare you so that you can make good design choices when facing a new learning problem!



Determine Type of Training Experience
Games against experts
Games against self
Table of correct moves
...
Determine Target Function
Board → move
Board → value
...
Determine Representation of Learned Function
Polynomial
Linear function of six features
Artificial neural network
...
Determine Learning Algorithm
Gradient descent
Linear programming
...
Completed Design