CSDS 440: Machine Learning

Soumya Ray (he/him, sray@case.edu)

Olin 516

Office hours T, Th 11:15-11:45 or by appointment

Announcements/Status

- Groups

- Cwru-courses/ Repositories

Play this 2-player game!

- Draw this board on a sheet of paper
- Each player chooses/makes three "pieces"
- Take turns placing your pieces on the intersections
- Once all pieces are placed, take turns to move them to any adjacent open intersection
- Goal: starting with empty board, first player to make three pieces in a line wins
- If no winner after 10 moves, draw

- Each person in a group should play at least 2 games (AB/AC/BC)
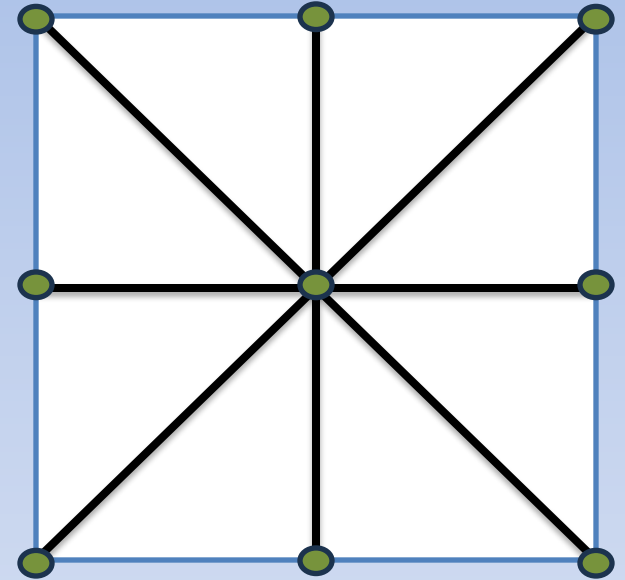
Today

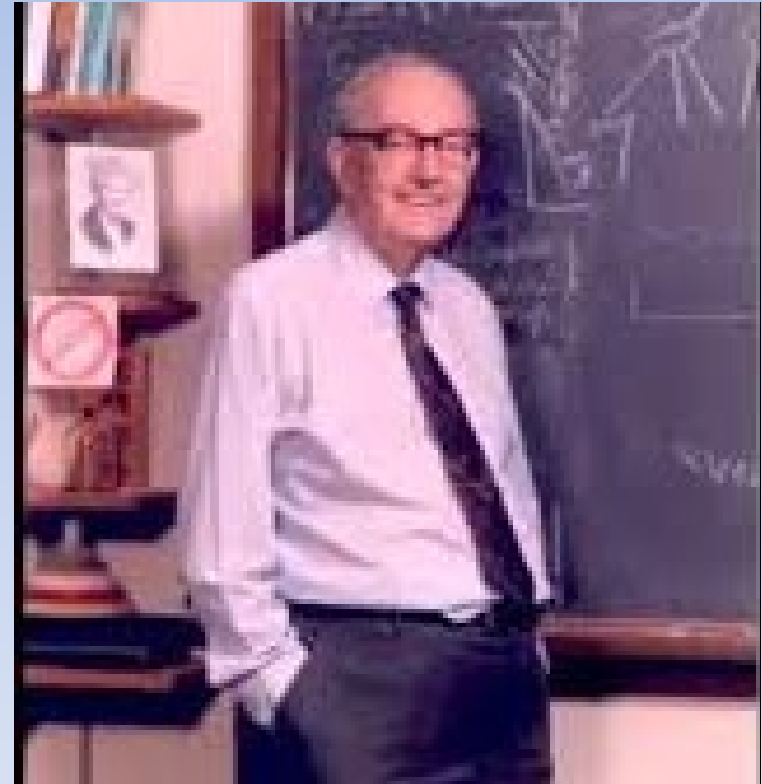- Foundations of Machine Learning

What is "machine learning"?

- "Machine"=autonomous system
  - No (or limited) human intervention
  - Robots, software agents, etc.

What is "Learning"?

"Learning denotes <u>changes</u> in the system that enable the system to do the same <u>task</u> more <u>effectively</u> the next time."

—Herbert Simon (1916-2001)

– Also, how to do *related* tasks more effectively

A Specification for a Learning System

- Given:
  - Learning task,
  - Task examples E,
  - Performance measure P

- Do: Produce a concept that is good with respect to P on *all* examples of the task
  - Measured by proxy on E

Example

- Learning Task: Learn to play a board game
- Performance measure?
  - Games won/lost
- Examples?
  - Games played (sequences of moves till win/loss)
- Concept?
  - Some function mapping current state of game to suitable moves to play

Idea

- If the learning system plays/sees enough games,

- And it produces a mapping from game state to moves (concept),

- And this concept does well with respect to the measure of "number of games won",

- Then the system has "learned to play the game"

Other Examples

- Learn to recognize lions
  - *E*: animals, annotated "lion" or "not-lion"
  - *P*: fraction of animals correctly recognized as lion/not-lion
- Learn to drive
  - *E*: sequence of road/traffic conditions and correct vehicle operation
  - *P*: distance traveled without accident

Two Phases of Learning

- "Learning" or "Training" phase
  - Reason about the examples E
  - Formulate a concept that does well w.r.t. P on E
  - Could also use any prior knowledge

- "Evaluation" or "Testing" phase
  - Use learned concept on future, novel examples

Online and Batch (Offline) Learning

- Batch/Offline Learning: one learning phase, with a large set of examples, followed by a testing phase

- Online learning: Examples arrive one at a time (or in small groups); learning and evaluation phases are iterated

Inductive Generalization

- In all learning problems, need to reason from specific examples to a general case
  - Memorization ≠ Learning

- Other kinds of reasoning
  - deduction (general to specific)
  - abduction (most likely cause)

Target Concept

- The unknown underlying concept that solves the learning task

  - E.g., "has-fur" and "long-teeth" and "looks-scary" $\rightarrow$ "lion"

- Typically, P will be a measure of difference between the learner's concept and the target concept, with respect to E

Hypothesis Space

- Defines the space of general concepts the learning system will consider
  - E.g., all possible conjunctions of animal properties
  - "has-fur" and "long-teeth" and "looks-scary", "has-fur" and "long-teeth" and NOT-"looks-scary" , "has-fur" and NOT-"long-teeth" and "looks-scary" ….

- Ideally, target concept is a member of this space
  - What if it isn't?
    - Maybe we should include *all possible* hypotheses?

No "Tabula Rasa" Learning

- A space that includes all possible hypotheses also
  - Contains many overly complex concepts
  - Contains the concept that memorizes *E*
    - Indistinguishable from target by any *P* (w.r.t. *E*)
  - *May be* too big to search feasibly

- For effective inductive generalization
  - **Must** restrict hypothesis space
  - while still (hopefully) keeping the target concept in it

Inductive Bias

- The set of assumptions used by a learning system to restrict its hypothesis space

- The more assumptions made, the "stronger" the bias

- Can quantify this (later)