Methods for Intelligent Systems

Lecture Notes on Machine Learning

Matteo Matteucci

matteucci@elet.polimi.it

Department of Electronics and Information

Politecnico di Milano

Some Course Info

- Lectures given by:
  - Matteo Matteucci (Lecturer) +39 02 2399 3470
    http://www.dei.polimi.it/people/matteucci
    matteucci@elet.polimi.it
  - Davide Eynard (Teaching Assistant) eynard@elet.polimi.it
    http://www.dei.polimi.it/people/eynard
  - Simone Tognetti (Teaching Assistant)
    tognetti@elet.polimi.it
    http://www.dei.polimi.it/people/tognetti
- Course Material on Machine Learning (tentative)
  - Lectures notes (teaching area of lecturer home page)
  - *Pattern Recognition 2nd Ed.*, R.O. Duda, P.E.Hart and D.G. Stork
  - *Pattern Recognition and Machine Learning*, C.M. Bishop.
  - *Machine Learning*, T. Mitchell.

More Important Course Info

- The Goals of this course:
  - Introduce Artificial Intelligence techniques to be applied in advanced (adaptive) applications
  - Provide a sound background on Machine Learning techniques
  - Introduce Data Mining basics

- Evaluation and Grading !!
  - Oral final examination theory and simple exercises [70% of grading]
  - A mandatory homework (due in 15 days) to integrate the theory; this can be substitued with a project to be discussed in advance with the teacher. [30% of grading]

What do we mean/expect with/from “Intelligent Systems”?

Why do we need “Intelligent Systems”?

Dartmouth 1955 – Conception of AI

A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence

J. McCarthy, Dartmouth College
M.L. Minsky, Harvard University
N. Rochester, I.B.M. Corporation
C.E. Shannon, Bell Telephone Laboratories

August 31, 1955

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

Artificial Intelligence: Do We Have a Single Definition?

- The branch of computer science that deal with writing computer programs that can solve problems creatively [*WordNet Dictionary*]
- The field of computer science concerned with the concepts and methods of symbolic inference by computer and symbolic knowledge representation for use in making inferences. AI can be seen as an attempt to model aspects of human thought on computers. It is also sometimes defined as trying to solve by computer any problem that a human can solve faster. [*Computing Dictionary*]
- The study of how to create a computer that can "think" like a human being. See also expert systems and machine learning. [*Biology Dictionary*]
- A field within computer science concerned with developing technology to enable computers to solve problems (or assist humans in solving problems) using explicit representations of knowledge and reasoning methods employing that knowledge. [*DARPA Document*]

Artificial Intelligence: Do We Have a Single Definition?

- The field of computer science concerned with the concepts and methods of symbolic inference by computer and symbolic knowledge representation for use in making inferences. AI can be seen as an attempt to model aspects of human thought on computers. It is also sometimes defined as trying to solve by computer any problem that a human can solve faster. [*The Free On-line Dictionary of Computing*]
- The simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using the rules to reach approximate or definite conclusions), and self-correction. Particular applications of AI include expert systems, speech recognition, and machine vision. [*searchCIO.com*]
- (1) the capability of a machine to imitate intelligent human behavior. (2) a branch of computer science dealing with the simulation of intelligent behavior in computers. [*Merriam Webster Dictionary*]

Artificial Intelligence: Yet Another Definition!

- Why intelligence has to be symbolic?
- Why should we consider as intelligence only human behavior?
- What's the difference between simulate and imitate/emulate?
- Is "faster than human" something related to Artificial Intelligence?
- How can we create "thinking" machines if we do not know what *thinking* is?

Artificial Intelligence: A field within computer science concerned with developing technology to enable computers to solve problems (or assist humans in solving problems) that according to common people could be solved only by humans.

"Making a machine behave in ways that would be called intelligent if a human were so behaving."

John McCarthy, 1955.

Dartmouth 1956 – The AI Program

1. Automatic Computers
2. How Can a Computer be Programmed to Use a Language
3. Neuron Nets
4. Theory of the Size of a Calculation
5. Self-Improvement
6. Abstractions
7. Randomness and Creativity

What about Self-Improvement?
Can you guess a definition of Machine Learning?

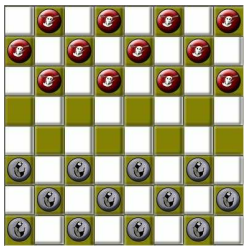
Machine Learning: the study or development of models and algorithms that make systems automatically improve their performance during execution.

Self-Improvement and Learning

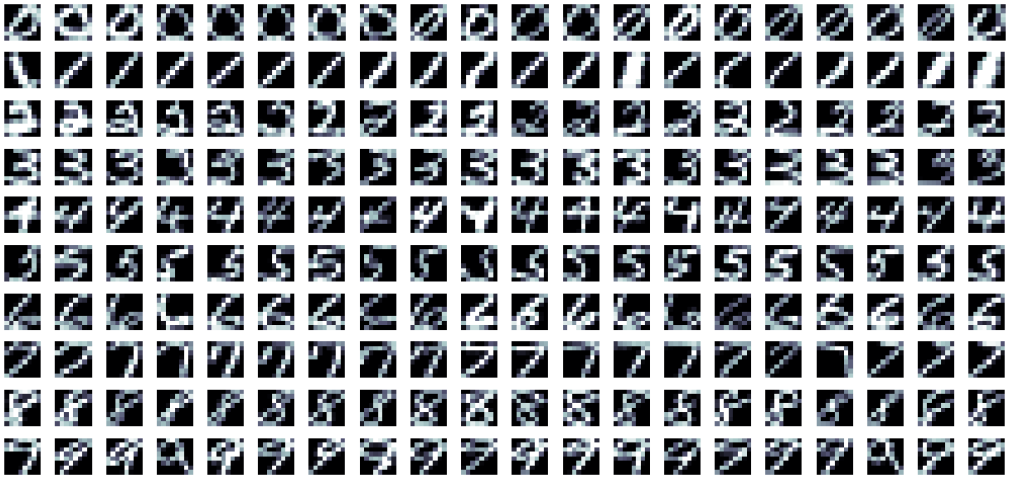
"A computer program is said to **learn** from experience **E** with respect to some class of **task T** and a **performance measure P**, if its performance at tasks in **T**, as measured by **P** improves because of experience **E**."

Tom Mitchell, 1997.

- Task **T**
  - Play checkers
- Experience **E**
  - Games played with other players
  - Games played against itself
- Performance **P**
  - Percentage of games won



Example: ZIP Codes Images



There are 7291 training observations and 2007 test observations.
Each observation is a 16 x 16 grayscale image

Learning Paradigms

Imagine an organism or machine that experiences a series of sensory inputs:

$$\mathbf{E} = x_1, x_2, x_3, x_4, \ldots$$

- *Supervised learning*: given the **desired outputs** $y_1, y_2, \ldots$, learn to produce the correct output given new input
- *Unsupervised learning*: **exploit regularities in** $\mathbf{E}$ to **build a representation** that can be used for reasoning or prediction
- *Reinforcement learning*: **producing actions** $a_1, a_2, \ldots$ which affect the environment, and **receiving rewards** $r_1, r_2, \ldots$ learn to act in a way that **maximises rewards** in the long term

Inductive Hypothesis: A solution that approximates the target function over a sufficiently large set of training examples will also approximate the target function over unobserved examples.

Applications: “What’s your flava?”

- Self customizing programs
  - Newsreader that learns user interests
  - Email anti-spam filters
  - User adaptive graphical interfaces
  - . . .
- Data mining
  - Extracting medical knowledge from medical records
  - Using historical data to improve decisions
  - . . .
- Software applications we can’t program by hand
  - Autonomous driving
  - Speech recognition
  - . . .

Do you have any question?

Course Outline [*Tentative*]

- Probability for Dataminers
  - Probability Basics & Information Gain
- Supervised Learning
  - Decision Trees & Decision Rules
  - Bayes Classifiers & Bayesian Networks
  - (Hidden) Markov Models
  - Instance Based Learning
  - Support Vector Machines
- Unsupervised Learning
  - Clustering
  - Data Reduction
- Model Selection Techniques
  - Cross-Validation & Model Complexity
  - Model Evaluation & Credibility