Machine learning

Introduction

Hamid Beigy

Sharif University of Technology

October 29, 2021

Table of contents

What is machine learning?

The field of *machine learning* is concerned with the question of how to construct computer programs that automatically improve with the experience.

Definition (Mohri et. al., 2012)

Computational methods that use experience to improve performance or to make accurate predictions.

Definition (Mitchel 1997)

A computer program is said to *learn*

- from *training experience E*
- with respect to some class of *tasks T*
- and *performance measure P*,

if its performance at tasks in T, as measured by P, improves with experience E.

Example (Checkers learning problem)

Class of task T: playing checkers.

Performance measure P: percent of games won against opponents.

Training experience E: playing practice game against itself.

Example (Handwriting recognition learning problem)

Class of 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.

Example (Robot driving learning problem)

Class of task T: driving a robot on the public highways using vision sensors.

Performance measure P: average distance travelled before an error.

Training experience E: a sequence of images and steering command recorded.

We need machine learning because

1. Tasks are too complex to program
   - Tasks performed by animals/humans such as driving, speech recognition, image understanding, and etc.
   - Tasks beyond human capabilities such as weather prediction, analysis of genomic data, web search engines, and etc.
2. Some tasks need adaptivity. When a program has been written down, it stays unchanged. In some tasks such as optical character recognition and speech recognition, we need the behavior to be adapted when new data arrives.

Types of machine learning

Machine learning algorithms based on the information provided to the learner can be classified into three main groups.

1. Supervised/predictive learning: The goal is to learn a mapping from inputs x to outputs y given the labeled set $S = \{(x_1, t_1), (x_2, t_2), \ldots, (x_N, t_N)\}$. x_k is called feature vector.
   - When $t_i \in \{1, 2, \ldots, C\}$, the learning problem is called classification.
   - When $t_i \in \mathbb{R}$, the problem is called regression.
2. Unsupervised/descriptive learning: The goal is to find interesting pattern in data $S = \{x_1, x_2, \ldots, x_N\}$. Unsupervised learning is arguably more typical of human and animal learning.
3. Reinforcement learning: Reinforcement learning is learning by interacting with an environment. A reinforcement learning agent learns from the consequences of its actions.

1. Supervised learning:
   - Classification:
     - Document classification and spam filtering.
     - Image classification and handwritten recognition.
     - Face detection and recognition.
   - Regression:
     - Predict stock market price.
     - Predict temperature of a location.
     - Predict the amount of PSA.
2. Unsupervised/descriptive learning:
   - Discovering clusters.
   - Discovering latent factors.
   - Discovering graph structures (correlation of variables).
   - Matrix completion (filling missing values).
   - Collaborative filtering.
   - Market-basket analysis (frequent item-set mining).
3. Reinforcement learning:
   - Game playing.
   - robot navigation.

The need for probability theory

- Probability theory can be applied to any problem involving uncertainty.
- A key concept in machine learning is uncertainty. In machine learning, uncertainty comes in many forms:
  - What is the best prediction about the future given some past data?
  - What is the best model to explain some data?
  - What measurement should I perform next?
- Data comes from a process that is not completely known.
- This lack of knowledge is indicated by modeling the process as a random process.
- The process actually may be deterministic, but we don't have access to complete knowledge about it, we model it as random and we use the probability theory to analyze it.
- The probabilistic approach to machine learning is closely related to the field of statistics, but differs slightly in terms of its emphasis and terminology[1].

[1] http://www-stat.stanford.edu/~tibs/stat315a/glossary.pdf

Outline of course

Outline of course

1. Introduction to machine learning & probability theory
2. Supervised learning:
    - Linear models for regression
    - Classifiers based on Bayes decision theory
    - Linear & Nonlinear models for classification
    - Combining classifiers
    - Evaluating classifiers
    - Computational learning theory
3. Unsupervised/descriptive learning:
    - Feature selection & Feature extraction/dimensionality reduction
    - Clustering & clustering evaluation
4. Reinforcement learning:
    - Reinforcement model & model-based learning
    - Monte-carlo & Temporal difference methods
5. Advanced topics:
    - Statistical learning theory
    - Graphical models
    - Deep & semi-supervised & Active & online learning
    - Large scale machine learning

Course Information

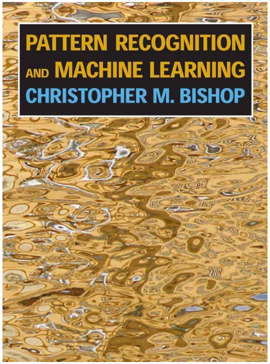
1. Course name : Machine learning
2. The objective of machine learning is the study of computer algorithms that can improve automatically through experience and by the use of data.
3. Instructor : Hamid Beigy
   Email : beigy@sharif.edu
4. Class Link: https://vc.sharif.edu/ch/beigy
5. Course Website: http://ce.sharif.edu/courses/00-01/1/ce717-1/
   http://sharif.edu/~beigy/14001-40717.html
6. Lectures: Sat-Mon (8:30-10:30)
7. TAs :
   Fariba Lotfi Email: flotfi@ce.sharif.edu
   Hassan Hamidi Email: hamidi3031@gmail.com

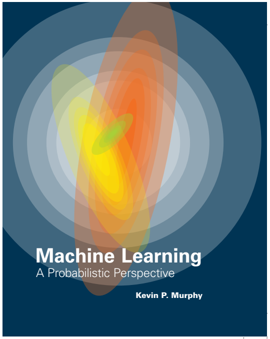
- Evaluation:

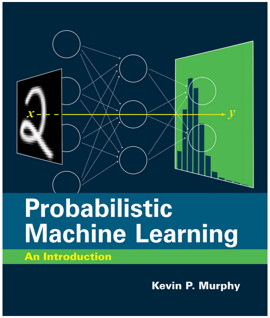
Mid-term exam	20%	1400-09-17	
Final exam	25%	1400-10-29	
Practical Assignments	35%		
Quiz	15%		
Paper	5%	1400-09-17	Hard deadline for selection

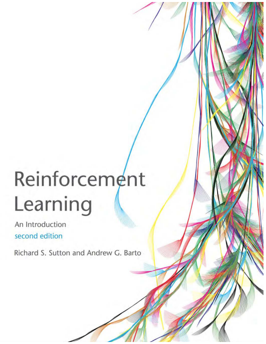
References

1. Main references









2. Other useful references are (Mitchell 1997), (Alpaydin 2014), (Hastie, Tibshirani, and Friedman 2009), and (Szepesvari 2010).

Relevant journals

1. IEEE Trans on Pattern Analysis and Machine Intelligence
2. Journal of Machine Learning Research
3. Pattern Recognition
4. Machine Learning
5. Neural Networks
6. Neural Computation
7. Neurocomputing
8. IEEE Trans. on Neural Networks and Learning Systems
9. Annuals of Statistics
10. Journal of the American Statistical Association
11. Pattern Recognition Letters
12. Artificial Intelligence
13. Data Mining and Knowledge Discovery
14. IEEE Transaction on Cybernetics (SMC-B)
15. IEEE Transaction on Knowledge and Data Engineering
16. Knowledge and Information Systems

Relevant conferences

1. Neural Information Processing Systems (NIPS)
2. International Conference on Machine Learning (ICML)
3. European Conference on Machine Learning (ECML)
4. Asian Conference on Machine Learning (ACML2013)
5. Conference on Learning Theory (COLT)
6. Algorithmic Learning Theory (ALT)
7. Conference on Uncertainty in Artificial Intelligence (UAI)
8. Practice of Knowledge Discovery in Databases (PKDD)
9. International Joint Conference on Artificial Intelligence (IJCAI)
10. IEEE International Conference on Data Mining series (ICDM)

1. Packages:
   - R http://www.r-project.org/
   - Weka http://www.cs.waikato.ac.nz/ml/weka/
   - RapidMiner http://rapidminer.com/
   - MOA http://moa.cs.waikato.ac.nz/
2. Datasets:
   - UCI Machine Learning Repository http://archive.ics.uci.edu/ml/
   - StatLib http://lib.stat.cmu.edu/datasets/
   - Delve http://www.cs.toronto.edu/~delve/data/datasets.html

1. Chapter 1 of Pattern Recognition and Machine Learning Book (Bishop 2006).
2. Chapter 1 of Machine Learning: A probabilistic perspective (Murphy 2012).
3. Chapter 1 of Probabilistic Machine Learning: An introduction (Murphy 2022).

- Alpaydin, Ethem (2014). *Introduction to Machine Learning*. 3rd ed. MIT Press.
- Bishop, Christopher M. (2006). *Pattern Recognition and Machine Learning*. Springer-Verlag.
- Hastie, Trevor, Robert Tibshirani, and Jerome Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. 2nd ed. Springer.
- Mitchell, Tom M. (1997). *Machine Learning*. McGraw-Hill.
- Murphy, Kevin P. (2012). *Machine Learning: A Probabilistic Perspective*. The MIT Press.
- – (2022). *Probabilistic Machine Learning: An introduction*. MIT Press.
- Szepesvari, Csaba (2010). *Algorithms for Reinforcement Learning*. Morgan and Claypool Publishers.

Questions?