Welcome to CMPS 142 and 242: Machine Learning

- Instructor: David Helmbold, dph@soe.ucsc.edu
- TA: Jiazhong Nie, niejiazhong@gmail.com
- Web page: www.soe.ucsc.edu/classes/cmps242/Spring11/
  - User: cmps242
  - Password: ceDQd2BW
- Text: Introduction to Machine Learning, by Alpaydin (2nd)

Administrivia

- Sign up sheet (enrollment)
- Evaluation: (142/242)
  - Homework 20 /20 %
  - Late midterm 50 / 30 %
  - Projects (group?) 30 / 50 %
- Expectations/Style
  - Reading assignments
  - Attendance/participation
  - My hearing/writing
  - Academic honesty

Topics:

- Introduction (ch1, 2)
- Bayesian learning and parameter estimation (ch 3-5)
- Instance based methods (ch 8)
- Decision Trees and Neural networks (ch 9, 11)
- Linear discrimination (ch 10)
- Support Vector Machines (ch 13)
- Boosting (AdaBoost) (ch 17)
- Clustering, EM Algorithm and K-means (ch 7)
- On-line prediction (Blum survey)

Introduction to Machine Learning

ETHEM ALPAYDIN
© The MIT Press, 2004
(modified by DPH 2006–2011)
alpaydin@boun.edu.tr
http://www.cmpe.boun.edu.tr/~ethem/i2ml

Why “Learn” ?

- Machine learning is programming computers to optimize a performance criterion using example data or past experience (inference in statistics)
- There is no need to “learn” to calculate payroll
- Learning is used when:
  - Human expertise does not exist (navigating on Mars),
  - Humans are unable to explain their expertise (speech recognition)
  - Solution changes in time (routing on a computer network)
  - Solution needs to be adapted or customized to particular cases (or users)

What We Talk About When We Talk About “Learning”

- Learning general models from a set of particular examples
- Data is cheap and abundant (data warehouses, data marts); knowledge is expensive and scarce.
- Example in retail: Customer transactions to consumer behavior:
  People who bought “Da Vinci Code” also bought “The Five People You Meet in Heaven” (www.amazon.com)
- Build a model that is a good and useful approximation to the data.
What is Machine Learning?

- Optimize a performance criterion using example data or past experience.
- Role of Statistics: Inference from a sample
- Role of Computer science: Efficient algorithms to
  - Solve the optimization problem
  - Representing and evaluating the model for inference

Stat. Machine learning is not:

- Cognitive science (how people think/learn)
- Teaching computers to think

But is related to:

- Statistics
- Data Mining - KDD
- Control theory
- part of AI, but not "traditional" AI

Data Mining Applications

- Retail: Market basket analysis, Customer relationship management (clustering)
- Finance: Credit scoring, fraud detection
- Manufacturing: Optimization, troubleshooting
- Medicine: Medical diagnosis
- Telecommunications: Quality of service optimization
- Bioinformatics: Motifs, alignment, protein structure
- Web mining: Search engines
- ...

Supervised Batch Learning

- Assume (unknown) distribution over "things"
- Get instances by drawing things from distribution and recording observations.
- Teacher labels instances making examples $$(x,r)$$
  - Or $$(x,y)$$ or $$(x,t)$$ or $$(x,r)$$
- Set of labeled examples is the training set or sample
- Create hypothesis (rule) from sample
- Hypothesis predicts on new random instances, scored by loss function

Learning Framework

- Classification: labels are nominal (unordered set, e.g. [ham, spam] {democrat, republican, indep.})
- Binary Classification
- Regression: labels are numeric (e.g. price of used car)
- Sometimes labels are probabilities

Supervised Learning (cont.)
Examples

<table>
<thead>
<tr>
<th>Thing</th>
<th>Observations</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written Digit</td>
<td>Pixel array</td>
<td>Which digit?</td>
</tr>
<tr>
<td>Email message</td>
<td>Words, Subject, sender</td>
<td>Ham or Spam?</td>
</tr>
<tr>
<td>Customer</td>
<td>Recent purchase</td>
<td>Interest level in a new product</td>
</tr>
<tr>
<td>Used car</td>
<td>Year, make, mpg, options</td>
<td>Price or value</td>
</tr>
</tbody>
</table>

Face Recognition

Training examples of a person

Test images

Regression

- Example: Price of a used car
- $x$: car attributes
- $y$: price
- $y = g(x | \theta)$
- $g()$ model (e.g. linear)
- $\theta$ parameters ($w$, $w_0$)

Supervised Learning: Uses

- Prediction of future cases: Use the rule to predict the output for future inputs
- Knowledge extraction: The rule is easy to understand
- Compression: The rule is simpler than the data it explains
- Outlier detection: Exceptions that are not covered by the rule, e.g., fraud and data entry errors

Other kinds of supervised learning

- Reinforcement learning - learning a policy for influencing or reacting to environment
  - No supervised output, but delayed rewards
  - Credit assignment problem
  - Game playing/robot in a maze, etc.
- On-line learning: predict on each instance in turn
- Semi-supervised learning uses both labeled and unlabeled data

Unsupervised Learning

- Learning “what normally happens”
- No labels
- Clustering: Grouping similar instances
- Example applications
  - Segmentation in customer relationship mgmt
  - Image compression: Color quantization
  - Bioinformatics: Learning motifs
  - Identifying unusual Airplane landings
Resources: Datasets

- Statlib: http://lib.stat.cmu.edu/
- Delve: http://www.cs.utoronto.ca/~delve/
- MLcomp: http://mlcomp.org

Resources: Journals

- Machine Learning
- Neural Computation
- Neural Networks
- IEEE Transactions on Neural Networks
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- Annals of Statistics
- Journal of the American Statistical Association
- ...

Resources: Conferences

- International Conference on Machine Learning (ICML)
  - ICML05: http://icml.ais.fraunhofer.de/
- European Conference on Machine Learning (ECML)
  - ECML05: http://ecmlpkdd05.ia.ac.cn/
- Neural Information Processing Systems (NIPS)
- Uncertainty in Artificial Intelligence (UAI)
  - UAI05: http://www.cs.toronto.edu/uai2005/
- Computational Learning Theory (COLT)
  - COLT05: http://learningtheory.org/colt2005/
- International Joint Conference on Artificial Intelligence (IJCAI)
  - IJCAI05: http://ijcai05.cs.ubc.ca/
- International Conference on Neural Networks (Europe)
  - ICANN05: http://www.inbspan.waw.pl/ICANN-2005/
- ...

Resources: Conferences