Machine Learning

Sargur N. Srihari
Machine Learning and AI

- If an expert system—brilliantly designed, engineered and implemented—cannot learn not to repeat its mistakes, it is not as intelligent as a worm or a sea anemone or a kitten.
  - Oliver G. Selfridge, from The Gardens of Learning.

- "Find a bug in a program, and fix it, and the program will work today. Show the program how to find and fix a bug, and the program will work forever."
  - Oliver G. Selfridge, in *AI's Greatest Trends and Controversies*
Machine Learning

- How to construct computer programs that automatically improve with experience
- Examples:
  - from medical records: treatments most effective
  - houses learn to optimize energy costs
  - search engines, computer games, adaptive user interfaces, personalized assistants, web bots, and scientific applications
Machine Learning

- Opens new uses of computers
  - new levels of competence and customization
  - better understanding of human learning abilities/disabilities
The Course

- Key algorithms and theory that form the core of machine learning
- Draws heavily from concepts and results from:
  - Statistics
  - Artificial Intelligence
  - Philosophy
  - Information theory
  - Biology
  - Computational complexity
  - Control theory
Books

Text Book:
- Tom Mitchell
  - Machine Learning, 1997
  - WCB McGraw Hill

Reference:
- Tony Jebara
  - Machine Learning: Discriminative and Generative, 2004
  - Kluwer
Student Background

- For undergraduate and graduate students in CSE, statistics and social sciences
- Two guidelines
  - Accessible to undergraduates
  - Material for PhD students to have before doing doctoral research in machine learning
The Material

- Balance of theory and practice
  - How does learning performance vary with number of samples
  - Which learning algorithms are appropriate for various learning tasks
Achievements in Machine Learning

- Recognizing spoken words
- Recognizing handwritten words
- Predict recovery rates for pneumonia patients
- Detect fraudulent use of credit cards
- Drive autonomous vehicles on highways
- Play backgammon approaching human world champion
Some Successful Applications of Machine Learning

- Learning to recognize spoken words
  - Speaker-specific strategies for recognizing primitive sounds (phonemes) and words from speech signal
  - Neural networks and methods for learning HMMs for customizing to individual speakers, vocabularies and microphone characteristics
Some Successful Applications of Machine Learning

- Learning to drive an autonomous vehicle
  - Train computer-controlled vehicles to steer correctly
  - Drive at 70 mph for 90 miles on public highways
  - Associate steering commands with image sequences
Problems Too Difficult To Program by Hand

ALVINN [Pomerleau] drives 70 mph on highways
Scientific Application of Machine Learning

- Learning to classify new astronomical structures
  - Very large databases to learn general regularities implicit in the data
  - Classify celestial objects from image data
  - Decision tree algorithms are now used by NASA to classify all objects in sky survey which consists of 3 terabytes of image data
Well-Posed Learning Problems

Definition:
A computer program is said to learn from experience \( E \)
with respect to some class of tasks \( T \) and performance measure \( P \),

if its performance at tasks \( T \), as measured by \( P \), improves with experience \( E \).
Well-defined Learning Problem

- Identify three features
  - class of tasks
  - measure of performance to be improved
  - source of experience
A Robot Driving Learning Problem

- **Task** $T$: driving on public, 4-lane highway using vision sensors
- **Performance measure** $P$: average distance traveled before an error (as judged by human overseer)
- **Training experience** $E$: a sequence of images and steering commands recorded while observing a human driver
A Handwriting Recognition Learning Problem

- **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
Handwriting Recognition Learning

<table>
<thead>
<tr>
<th>alarm</th>
<th>clock</th>
<th>did</th>
<th>not</th>
</tr>
</thead>
<tbody>
<tr>
<td>alarm</td>
<td>code</td>
<td>soil</td>
<td>rout</td>
</tr>
<tr>
<td>circle</td>
<td>raid</td>
<td>hot</td>
<td></td>
</tr>
<tr>
<td>shute</td>
<td>risk</td>
<td></td>
<td>not</td>
</tr>
<tr>
<td>clock</td>
<td>visit</td>
<td>riot</td>
<td></td>
</tr>
<tr>
<td>did</td>
<td></td>
<td>most</td>
<td></td>
</tr>
</tbody>
</table>

Wake me up this morning.

thai
taxis
tie
tier
this
moving
having
morning
running
loving
Text Categorization Problem

- **Task** $T$: assign a document to its content category
- **Performance measure** $P$: Precision and Recall
- **Training experience** $E$: Example pre-classified documents
Broad Definition of Learning

- Include most tasks conventionally called "learning"
- Encompass computer programs that improve from experience in straightforward ways:
  - DBMS that allows users to update entries which improves answering database queries
- Goal is to:
  - Define precisely a class of problems that forms interesting forms of learning, explore algorithms to solve such problems, understand fundamental structure of learning problems and processes
Disciplines: Influence on Machine Learning

- Artificial intelligence
  - learning symbolic representations of concepts
  - machine learning as a search problem
  - learning as an approach to improving problem solving

- Pattern recognition
  - Bayes’ theorem for calculating hypothesis probabilities
  - Naïve Bayes classifier
Disciplines: Influence on Machine Learning

- Computational complexity theory
  - Theoretical bounds on complexity of different learning tasks measured in terms of no of training samples, no of mistakes
Disciplines: Influence on Machine Learning

- Control theory
  - procedures to control processes to optimize predefined objectives and predict next state of process controlled

- Information theory
  - measures of entropy, information content
  - minimum description length approaches to learning
  - optimal codes and relationship to optimal training sequences for encoding hypothesis
Disciplines: Influence on Machine Learning

- Philosophy
  - Occam’s razor--simplest hypothesis is best (decision trees)
  - justifying generalizing beyond observed data
Disciplines: Influence on Machine Learning

- Psychology and neurobiology
  - power law of practice-- human response improves with performance

- Statistics
  - characterization of errors (eg, bias and variance) that occur when estimating the accuracy of a hypothesis based on a limited sample of data
  - confidence intervals, statistical tests
Perspectives & Issues

- Machine Learning involves searching a very large space of possible hypotheses that fits observed data and any prior knowledge held by the observer.
Issues in Machine Learning

- What algorithms exist?
- How much training data is sufficient?
- When and how can prior knowledge held by the learner guide the process of generalizing from examples?
Issues in Machine Learning

- What is the best strategy for choosing a useful next training experience?
- What is the best way to reduce the learning task to one or more function approximation systems?
- How can the learner automatically alter its representation to represent and learn the target function?
Topics (Mitchell Text)

- Concept Learning: general to specific ordering of hypotheses
- Decision Tree Learning, Occam’s razor
- Artificial Neural Networks, Backpropagation Algorithm
- Statistics and Estimation Theory in evaluating hypotheses
Topics Covered

- Naïve Bayes Classifier
- Instance based learning: nearest neighbor learning
- Learning algorithms modeled after biological evolution: Genetic algorithms and Genetic programming
Summary

- Machine learning algorithms have great practical value in a variety of application domains
- Machine learning draws on ideas from a diverse set of disciplines
- A well-defined learning problem requires a well-specified task, performance metric, and source of experience
- Designing a machine learning approach involves a number of design choices
- Learning involves search